

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 September 2003 (04.09.2003)

PCT

(10) International Publication Number
WO 03/072041 A2

(51) International Patent Classification⁷: **A61K**

NJ 07065-0907 (US). **SIMON, Adam, J.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). **ZUCK, Paul, D.** [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

(21) International Application Number: **PCT/US03/05458**

(74) Common Representative: **MERCK & CO., INC.**; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

</div

TITLE OF THE INVENTION**ASSAYS TO MONITOR AMYLOID PRECURSOR PROTEIN PROCESSING****CROSS-REFERENCE TO RELATED APPLICATIONS**

5 This application claims the benefit of U.S. Provisional Application No. 60/360,274, filed February 27, 2002, the contents of which are incorporated herein by reference in their entirety.

STATEMENT REGARDING FEDERALLY-SPONSORED R&D

10 Not applicable.

REFERENCE TO MICROFICHE APPENDIX

 Not applicable.

15 **FIELD OF THE INVENTION**

 The present invention is directed to the field of Alzheimer's disease. In particular, the present invention provides novel methods of identifying substances that are specific inhibitors of various steps in the processing of amyloid precursor protein.

20

BACKGROUND OF THE INVENTION

 Alzheimer's disease is a common, chronic neurodegenerative disease, characterized by a progressive loss of memory and sometimes severe behavioral abnormalities, as well as an impairment of other cognitive functions that often leads to dementia and death. It ranks as the fourth leading cause of death in industrialized societies after heart disease, cancer, and stroke. The incidence of Alzheimer's disease is high, with an estimated 2.5 to 4 million patients affected in the United States and perhaps 17 to 25 million worldwide. Moreover, the number of sufferers is expected to grow as the population ages.

25

 A characteristic feature of Alzheimer's disease is the presence of large numbers of insoluble deposits, known as amyloid plaques, in the brains of those affected. Autopsies have shown that amyloid plaques are found in the brains of virtually all Alzheimer's patients and that the degree of amyloid plaque deposition correlates with the degree of dementia (Cummings & Cotman, 1995, Lancet

326:1524-1587). While some opinion holds that amyloid plaques are a late stage by-product of the disease process, the consensus view is that amyloid plaques are more likely to be intimately, and perhaps causally, involved in Alzheimer's disease.

A variety of experimental evidence supports this view. For example,

5 A β , a primary component of amyloid plaques, is toxic to neurons in culture and transgenic mice that overproduce A β in their brains show significant deposition of A β into amyloid plaques as well as significant neuronal toxicity (Yankner, 1990, Science 250:279-282; Mattson et al., 1992, J. Neurosci. 12:379-389; Games et al., 1995, Nature 373:523-527; LaFerla et al., 1995, Nature Genetics 9:21-29). Mutations in the

10 APP gene, leading to increased A β production, have been linked to heritable forms of Alzheimer's disease (Goate et al., 1991, Nature 349:704-706; Chartier-Harlan et al., 1991, Nature 353:844-846; Murrell et al., 1991, Science 254:97-99; Mullan et al., 1992, Nature Genetics 1:345-347). Presenilin-1 (PS1) and presenilin-2 (PS2) related familial early-onset Alzheimer's disease (FAD) shows disproportionately increased

15 production of A β 1-42, the 42 amino acid isoform of A β , as opposed to A β 1-40, the 40 amino acid isoform (Scheuner et al., 1996, Nature Medicine 2:864-870). The longer isoform of A β is more prone to aggregation than the shorter isoform (Jarrett et al., 1993, Biochemistry 32:4693-4697). Injection of the insoluble, fibrillar form of A β into monkey brains results in the development of pathology (neuronal destruction, tau

20 phosphorylation, microglial proliferation) that closely mimics Alzheimer's disease in humans (Geula et al., 1998, Nature Medicine 4:827-831). See Selkoe, 1994, J. Neuropathol. Exp. Neurol. 53:438-447 for a review of the evidence that amyloid plaques have a central role in Alzheimer's disease.

A β , a 39-43 amino acid peptide derived by proteolytic cleavage of the

25 amyloid precursor protein (APP), is the major component of amyloid plaques (Glenner & Wong, 1984, Biochem. Biophys. Res. Comm. 120:885-890). APP is actually a family of polypeptides produced by alternative splicing from a single gene. Major forms of APP are known as APP695, APP751, and APP770, with the

30 subscripts referring to the number of amino acids in each splice variant (Ponte et al., 1988, Nature 331:525-527; Tanzi et al., 1988, Nature 331:528-530; Kitaguchi et al., 1988, Nature 331:530-532). APP is membrane bound and undergoes proteolytic cleavage by at least two pathways. In one pathway, cleavage by an enzyme known as α -secretase occurs while APP is still in the trans-Golgi secretory compartment (Kuentzel et al., 1993, Biochem. J. 295:367-378). This cleavage by α -secretase

occurs within the A β portion of APP, thus precluding the formation of A β . In another proteolytic pathway, cleavage of the Met596-Asp597 bond (numbered according to the 695 amino acid protein) by an enzyme known as β -secretase occurs. This cleavage by β -secretase generates the N-terminus of A β . The C-terminus is formed
5 by cleavage by a second enzyme known as γ -secretase. The C-terminus is actually a heterogeneous collection of cleavage sites rather than a single site since γ -secretase activity occurs over a short stretch of APP amino acids rather than at a single peptide bond. Peptides of 40 or 42 amino acids in length (A β 1-40 and A β 1-42, respectively) predominate among the C-termini generated by γ -secretase. A β 1-42 is more prone to
10 aggregation than A β 1-40, is the major component of amyloid plaque (Jarrett et al., 1993, Biochemistry 32:4693-4697; Kuo et al., 1996, J. Biol. Chem. 271:4077-4081), and its production is closely associated with the development of Alzheimer's disease (Sinha & Lieberburg, 1999, Proc. Natl. Acad. Sci. USA 96:11049-11053). The bond cleaved by γ -secretase appears to be situated within the transmembrane domain of
15 APP. It is unclear as to whether the C-termini of A β 1-40 and A β 1-42 are generated by a single γ -secretase protease with sloppy specificity or by two distinct proteases. For a review that discusses APP and its processing, see Selkoe, 1998, Trends Cell. Biol. 8:447-453.

Much interest has focused on the possibility of inhibiting the
20 development of amyloid plaques as a means of preventing or ameliorating the symptoms of Alzheimer's disease. To that end, a promising strategy is to inhibit the activity of β - and γ -secretase, the two enzymes that together are responsible for producing A β . This strategy is attractive because, if the formation of amyloid plaques as a result of the deposition of A β is a cause of Alzheimer's disease, inhibiting the
25 activity of one or both of the two secretases would intervene in the disease process at an early stage, before late-stage events such as inflammation or apoptosis occur. Such early stage intervention is expected to be particularly beneficial (see, e.g., Citron, 2000, Molecular Medicine Today 6:392-397).

To that end, various assays have been developed that are directed to the
30 identification of compounds that may interfere with the production of A β or its deposition into amyloid plaques. U.S. Patent No. 5,441,870 is directed to methods of monitoring the processing of APP by detecting the production of amino terminal fragments of APP. U.S. Patent No. 5,605,811 is directed to methods of identifying inhibitors of the production of amino terminal fragments of APP. U.S. Patent No.

5,593,846 is directed to methods of detecting soluble A β by the use of binding substances such as antibodies. Esler et al., 1997, Nature Biotechnology 15:258-263 described an assay that monitored the deposition of A β from solution onto a synthetic analogue of an amyloid plaque. The assay was suitable for identifying compounds
5 that could inhibit the deposition of A β . However, this assay is not suitable for identifying substances, such as inhibitors of β - or γ -secretase, that would prevent the formation of A β .

Various groups have cloned and sequenced cDNA encoding a protein that is believed to be β -secretase (Vassar et al., 1999, Science 286:735-741; Hussain 10 et al., 1999, Mol. Cell. Neurosci. 14:419-427; Yan et al., 1999, Nature 402:533-537; Sinha et al., 1999, Nature 402:537-540; Lin et al., 2000, Proc. Natl. Acad. Sci. USA 97:1456-1460) but the identity of γ -secretase has been more elusive. A pair of proteins known as presenilin-1 and presenilin-2 are viewed as possible candidates (Selkoe & Wolfe, 2000, Proc. Natl. Acad. Sci. USA 97:5690-5692).

15 Presenilin-1 (PS1) and presenilin-2 (PS2) are polytopic membrane proteins that are involved in γ -secretase-mediated processing of APP. The most common cause of familial early-onset Alzheimer's disease is the autosomal dominant inheritance of assorted mutations in the PS1 gene (Sherrington et al., 1995, Nature 375:754-760). These PS1 mutations lead to increased production of A β 1-42
20 (Scheuner et al., 1996, Nature Medicine 2:864-870; Duff et al., 1996, Nature 383:710-713; Borchelt et al., 1996, Neuron 17:1005-1013). Similarly, certain mutations in PS2 cause familial early-onset Alzheimer's disease and increased generation of A β 42 (Levy-Lahad et al., 1995, Science 269:970-973). Cultured isolated neurons from PS1-deficient mice exhibit reduced γ -secretase-mediated
25 cleavage of APP (De Strooper et al., 1998, Nature 391:387-390). It was suggested that PS1 might influence trafficking of APP and/or γ -secretase or it might play a more direct role in proteolytic cleavage of APP. Directed mutagenesis of two conserved transmembrane-situated aspartates in PS1 was shown to inactivate γ -secretase activity in cellular assays, suggesting that PS1 is either a required diaspartyl cofactor for γ -
30 secretase or is itself γ -secretase (Wolfe et al., 1999, Nature 398:513-517). Moreover, Li et al., 2000, Nature 405:689-694 made photoactivatable derivatives of a highly specific and potent aspartyl protease transition state analog inhibitor and found that the inhibitor selectively labeled presenilin fragments.

Co-immunoprecipitation experiments have shown that PS1 and PS2 interact directly with the immature forms of APP in the endoplasmic reticulum where the disease-associated amyloid A β 1-42 peptide is probably generated (Xia et al., 1997 Proc. Natl. Acad. Sci. USA 94:8208-8213; Weidemann et al., 1997, Nat. Med. 3:328-332). Knock-out of PS1 activity greatly diminishes γ -secretase cleavage of APP (De Strooper et al., 1998, Nature 391:387-390). PS1 knock-outs do not exhibit total lack of γ -secretase activity but knock-out of both PS1 and PS2 activity does result in a total loss of γ -secretase activity (Herreman et al., 2000, Nat. Cell. Biol. 2:461-462; Zhang et al., 2000, Nat. Cell Biol. 2:463-465), suggesting that PS2 has a similar function to PS1 in the processing of APP.

Karlström et al., (Journal of Biological Chemistry papers in press, published on December 13, 2001 as Manuscript C100649200) describes an assay designed specifically to identify inhibitors of γ -secretase cleavage of APP. The authors inserted the GAL4 DNA binding domain fused to the VP16 transactivation domain into C99, a portion of APP containing the 99 carboxy-terminal amino acids. This fragment of APP contains the γ -secretase cleavage site but lacks the β -secretase cleavage site. Transaction of a UAS reporter plasmid by GAL4-VP16 confirmed cleavage of the Gal4-VP16/C99 substrate by γ -secretase only. Thus, the assay is capable of detecting γ -secretase inhibitors but not inhibitors of β -secretase or other modulators of APP processing requiring the N-terminal domain of APP.

Cao & Südhoff, 2001, Science 293:115-120 described work in which the GAL4 and LexA DNA binding domains were inserted into APP to demonstrate the potential of the cleaved C-terminus of APP for transcriptional co-activation. In this article, a transcriptional factor was not fused to APP and no attempt was made to develop an assay for the identification of APP processing inhibitors.

Sisodia, 1992, Proc. Natl. Acad. Sci. USA 89:6975-6979 described various changes in the amino acid sequence of APP in the region of the α -secretase cleavage site and the effect of those changes on cleavage by α -secretase. A change of K to V at position 612 of the 695 amino acid version of APP led to reduced cleavage by α -secretase.

U.S. Patent No. 6,333,167 B1 discloses an assay involving DNA constructs encoding portions of membrane proteins containing sites that are susceptible to cleavage by proteases that are fused to transcriptional repressors. Such constructs are introduced into cells that contain a reporter gene under the control of a

promoter that is sensitive to the repressor. In the absence of an inhibitor of the protease, the fusion protein is cleaved by the protease, releasing a membrane protein/repressor fusion protein that translocates to the nucleus and represses transcription from the reporter gene. In the presence of an inhibitor of the protease,
5 the membrane protein/repressor fusion protein is not released and thus cannot repress transcription from the reporter. An increase in reporter expression can therefore be used as a readout for the presence of an inhibitor.

SUMMARY OF THE INVENTION

10 The present invention is directed to methods of identifying inhibitors of the processing of amyloid precursor protein (APP) that are capable of identifying inhibitors of a number of steps of such processing. Unlike prior methods, the methods of the present invention can be used to screen for inhibitors of β -secretase cleavage, γ -secretase cleavage, APP extracellular signaling, or APP cytoplasmic signaling in a
15 single assay.

The methods employ a recombinant eukaryotic cell that is capable of processing APP. The cell has been engineered to express a fusion protein that contains amino acid sequences encompassing both the β -secretase cleavage site of APP and the γ -secretase cleavage site. The fusion protein also contains a transcription
20 factor fused in frame to the APP sequences.

When the recombinant cell is further engineered to contain a reporter gene, in which transcription of the reporter gene is driven by a regulatory DNA sequence that is inactive in the absence of the transcription factor but active in the transcription factor's presence, a system useful for screening for APP processing
25 inhibitors is provided. Since the recombinant cell has been selected so as to be capable of processing APP, the fusion protein will be processed, releasing the transcription factor and activating transcription of the reporter gene. The reporter gene has been preselected so that activation of the reporter gene leads to a detectable phenotype.

30 The system is utilized by exposing the recombinant cell to substances that are to be tested for the ability to inhibit APP processing. Those substances that are actually inhibitors of APP processing will cause diminished processing of the fusion protein, leading to smaller amounts of the transcription factor being released.

This leads to less transcription of the reporter gene. This results in a decrease in the phenotypic effect of the reporter gene that can be observed.

BRIEF DESCRIPTION OF THE DRAWINGS

5 Figure 1A-G shows a schematic diagram of several APP/transcription factor fusion constructs.

Figure 2A-B shows the DNA sequence (SEQ ID NO:1) of the fusion protein APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695).

10 Figure 3 shows the amino acid sequence (SEQ ID NO:2) of the fusion protein APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = K612V mutation; 4 = region of γ -secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

15 Figure 4A-B shows the DNA sequence (SEQ ID NO:3) of the fusion protein APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695).

20 Figure 5 shows the amino acid sequence (SEQ ID NO:4) of the fusion protein APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = K612V mutation; 4 = region of γ -secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

25 Figure 6A-C shows the DNA sequence (SEQ ID NO:5) of the fusion protein APP(1-651)SW, K612V, GAL4-VP16(delMet) APP (664-695).

30 Figure 7 shows the amino acid sequence (SEQ ID NO:6) of the fusion protein APP(1-651)SW, K612V, GAL4-VP16(delMet) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = K612V mutation; 4 = region of γ -secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

35 Figure 8A-C shows the DNA sequence (SEQ ID NO:7) of the fusion protein APP(1-651)wt, K612V, GAL4-VP16(del Met) APP (664-695).

40 Figure 9 shows the amino acid sequence (SEQ ID NO:8) of the fusion protein APP(1-651)wt, K612V, GAL4-VP16(del Met) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = K612V mutation; 4 = region of γ -

secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 10A-B shows the DNA sequence (SEQ ID NO:9) of the fusion protein APP(1-651)SW, TATexonI(M1L) APP (664-695).

5 Figure 11 shows the amino acid sequence (SEQ ID NO:10) of the fusion protein APP(1-651)SW, TATexonI(M1L) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = wild-type K at position 612; 4 = region of γ -secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of
10 APP.

Figure 12A-B shows the DNA sequence (SEQ ID NO:11) of the fusion protein APP(1-651)wt, TATexonI(M1L) APP (664-695).

15 Figure 13 shows the amino acid sequence (SEQ ID NO:12) of the fusion protein APP(1-651)wt, TATexonI(M1L) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = wild-type K at position 612; 4 = region of γ -secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of
APP.

20 Figure 14A-C shows the DNA sequence (SEQ ID NO:13) of the fusion protein APP(1-651)SW, GAL4-VP16(delMet) APP (664-695).

Figure 15 shows the amino acid sequence (SEQ ID NO:14) of the fusion protein APP(1-651)SW, GAL4-VP16(delMet) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = wild-type K at position 612; 4 = region of γ -
25 secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 16A-C shows the DNA sequence (SEQ ID NO:15) of the fusion protein APP(1-651)wt, GAL4-VP16(delMet) APP (664-695).

30 Figure 17 shows the amino acid sequence (SEQ ID NO:16) of the fusion protein APP(1-651)wt, GAL4-VP16(delMet) APP (664-695) with the various different regions of the fusion protein demarcated. 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = wild-type K at position 612; 4 = region of γ -secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 18A-B shows the cDNA sequence (SEQ ID NO:17) and Figure 18C shows the amino acid sequence (SEQ ID NO:18) of the 695 amino acid splice variant of wild-type Alzheimer's precursor protein (APP). See GenBank accession no. Y00264 and Kang et al., 1987, Nature 325:733-736.

5 Figure 19 shows data from an embodiment in which the assay of the present invention was used to identify both a β -secretase inhibitor and a γ -secretase inhibitor. See Example 3 for details.

Figure 20 shows a schematic diagram of pCR2.1 Gal4-VP16.

10 Figure 21A shows a schematic diagram of pRBR121. Figure 21B shows a schematic diagram of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695).

15 Figure 22A shows a schematic diagram of pRBR186. Figure 22B shows a schematic diagram of the viral plasmid pNL4-3. Figure 22C shows a schematic diagram of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) with additional details as compared to Figure 21B, which shows the same plasmid.

Figure 23 shows a schematic diagram of pRSV Kan/Neo res.

Figure 24 shows a schematic diagram of pUCd5TAT.

20 Figure 25A shows a schematic diagram of pMM321. Figure 25B-D shows the nucleotide sequence of pMM321. The upper strand is SEQ ID NO:19. The lower strand (SEQ ID NO:20) is the reverse complement of SEQ ID NO:19.

25 Figure 26A shows a schematic diagram of the expression vector pcDNA3.1 zeo (+)APP(1-651)SW, K612V-(M1L)TATexonI. This expression vector directs the expression of a fusion protein containing the first 651 amino acids of APP with the Swedish version of the β -secretase cleavage site and the K612V mutation fused to the first exon of HIV1 TAT. The methionine at position 1 of TAT has been changed to leucine. Figure 26B-G shows the nucleotide sequence of pcDNA3.1 zeo (+)APP(1-651)SW, K612V-(M1L)TATexonI. The upper strand is SEQ ID NO:21. The lower strand (SEQ ID NO:22) is the reverse complement of SEQ ID NO:21.

30 Figure 27A-B shows a schematic diagram depicting general features of the present invention. Figure 27A: The vertical bar represents a fusion protein with APP sequences represented as unfilled or lightly shaded portions of the bar. The lightly shaded portion represents A β . "BACE" indicates the β -secretase cleavage site. The dark shaded portion represents the transcription factor fused between APP

sequences. The horizontal bar represents a membrane in which the uncleaved fusion protein is embedded, e.g., the endoplasmic reticulum. Figure 27B: The transcription factor (plus small amounts of APP), having been released from the fusion protein and thus the membrane by APP processing, is shown in the nucleus binding to and 5 activating the regulatory DNA sequence ("Transcription Factor Response Element") that controls the expression of the reporter gene.

Figure 28A-B shows the DNA sequence (SEQ ID NO:23) of the fusion protein APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695).

Figure 29 shows the amino acid sequence of a fusion protein (APP(1-10 651)NFEV, K612V-TATexonI(M1L) APP (664-695)) (SEQ ID NO:24) containing the sequence NFEV at the β -secretase cleavage site (underlined at 2). The other portions of the fusion protein are indicated as follows: 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = K612V mutation; 4 = region of γ -secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

15 Figure 30A-C shows the DNA sequence (SEQ ID NO:25) of the fusion protein APP(1-651)NFEV, K612V, GAL4-VP16(delMet) APP (664-695).

Figure 31 shows the amino acid sequence of a fusion protein (APP(1-651)NFEV, K612V, GAL4-VP16(delMet) APP (664-695)) (SEQ ID NO:26) containing the sequence NFEV at the β -secretase cleavage site (underlined at 2). The 20 other portions of the fusion protein are indicated as follows: 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 = K612V mutation; 4 = region of γ -secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 32A shows a schematic diagram of pcDNA3.1 zeo (+), a 25 eukaryotic expression vector that is suitable for use in the present invention. Figure 32B-F shows the nucleotide sequence of pcDNA3.1 zeo (+). The upper strand is SEQ ID NO:27. The lower strand (SEQ ID NO:28) is the reverse complement of SEQ ID NO:27.

Figure 33 shows data from an embodiment of the present invention 30 utilizing a β -galactosidase reporter gene in which the assay of the present invention was used to identify both a β -secretase inhibitor and a γ -secretase inhibitor. See Example 8 for details.

Figure 34 shows data from an embodiment of the present invention in which a fusion protein having a wild-type β -secretase cleavage site and a fusion

protein having a Swedish β -secretase cleavage site are compared. See Example 9 for details.

Figure 35A-B shows the DNA sequence (SEQ ID NO:29) of the fusion protein APP(1-651)NFEV, TATexonI(M1L) APP (664-695).

5 Figure 36 shows the amino acid sequence of a fusion protein (APP(1-651)NFEV, TATexonI(M1L) APP (664-695)) (SEQ ID NO:30) containing the sequence NFEV at the β -secretase cleavage site (underlined at 2) and a wild-type K at position 612 (underlined at 3). The other portions of the fusion protein are indicated as follows: 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 =
10 wild-type K; 4 = region of γ -secretase cleavage; 5 = linker; 6 = TAT exon I; 7 = linker; 8 = amino acids 664-695 of APP.

Figure 37A-C shows the DNA sequence (SEQ ID NO:31) of the fusion protein APP(1-651)NFEV, GAL4-VP16(delMet) APP (664-695).

15 Figure 38 shows the amino acid sequence of a fusion protein (APP(1-651)NFEV, GAL4-VP16(delMet) APP (664-695)) (SEQ ID NO:32) containing the sequence NFEV at the β -secretase cleavage site (underlined at 2) and a wild-type K at position 612 (underlined at 3). The other portions of the fusion protein are indicated as follows: 1 = amino acids 1-651 of APP; 2 = region of β -secretase cleavage; 3 =
20 wild-type K; 4 = region of γ -secretase cleavage; 5 = linker; 6 = GAL4-VP16; 7 = linker; 8 = amino acids 664-695 of APP.

DETAILED DESCRIPTION OF THE INVENTION

For the purposes of this invention:

A “fusion protein” is a protein that contains at least two polypeptide regions and, optionally, a linking peptide to operatively link the two polypeptides into one continuous polypeptide. The at least two polypeptide regions in a fusion protein are derived from different sources, and therefore a fusion protein comprises two polypeptide regions not normally joined together in nature.

30 A “linking sequence (or linker peptide)” contains one or more amino acid residues joined in peptide bonds. A linking sequence serves to join two polypeptide regions of differing origins in a fusion protein via a peptide bond between the linking sequence and each of the polypeptide regions.

Typically, a fusion protein is synthesized as a continuous polypeptide in a recombinant host cell which contains an expression vector comprising a

nucleotide sequence encoding the fusion protein where the different regions of the fusion protein are fused in frame on either side of a linker peptide's coding sequence. The chimeric coding sequence (encoding the fusion protein) is operatively linked to expression control sequences (generally provided by the expression vector) that are functional in the recombinant host cell.

5 "Reporter gene," as used in the present invention, does not mean a DNA sequence present on the chromosome of a cell, generally possessing introns, as is often meant by the word "gene" in the art. Rather "reporter gene" means any DNA sequence encoding a protein or polypeptide that can give rise to a signal that can be detected or measured. "Reporter gene" does not mean a portion of the amino acid sequence of APP. "Reporter gene" will usually mean a DNA sequence, generally a cDNA sequence (although in some cases a reporter gene may have introns) that encodes a protein or polypeptide that is commonly used in the art to provide a measurable phenotype that can be distinguished over background signals.

10 15 A "nuclear localization signal (NLS)" is a region of a polypeptide which targets the polypeptide to the nucleus of the cell. One such NLS is that from the SV40 large T antigen. See, e.g., U.S. Patent No. 5,589,392; Kalderon et al., 1984, Cell 39:499-509. The minimum region of the SV40 large T antigen with NLS activity is Pro-Lys-Lys-Lys-Arg-Lys-Val (SEQ ID NO:22). See also U.S. Patent No.

20 5,776,689.

25 "Substances" that are screened in the present invention can be any substances that are generally screened in the pharmaceutical industry during the drug development process. For example, substances may be low molecular weight organic compounds (*e.g.*, having a molecular weight of less than about 2,000 daltons and preferably less than about 1,000 daltons), RNA, DNA, antibodies, peptides, or proteins. Substances are often tested in the methods of the present invention as large collections of substances, *e.g.* libraries of low molecular weight organic compounds, peptides, or natural products.

30 The conditions under which substances are employed in the methods described herein are conditions that are typically used in the art for the study of protein-ligand interactions or enzyme inhibition studies: *e.g.*, salt conditions such as those represented by such commonly used buffers as PBS or in tissue culture media; a temperature of about 4°C to about 55°C; incubation times of from several seconds to several hours or even up to 24 or 48 hours. Screening for the identification of

enzyme-specific inhibitors is a well-known procedure in the pharmaceutical arts and the numerous conditions under which such screening has been done are available in the literature to guide the practitioner of the present invention.

A "conservative amino acid substitution" refers to the replacement of one amino acid residue by another, chemically similar, amino acid residue. Examples of such conservative substitutions are: substitution of one hydrophobic residue (isoleucine, leucine, valine, or methionine) for another; substitution of one polar residue for another polar residue of the same charge (*e.g.*, arginine for lysine; glutamic acid for aspartic acid); substitution of one aromatic amino acid (tryptophan, tyrosine, or phenylalanine) for another.

"Transfection" refers to any of the methods known in the art for introducing DNA into a cell, *e.g.*, calcium phosphate or calcium chloride mediated transfection, electroporation, infection with a retroviral vector.

The present invention relates to the discovery of an assay system that permits the simultaneous screening for inhibitors of several types of amyloid precursor protein (APP) processing or signaling (*e.g.*, β -secretase cleavage, γ -secretase cleavage, APP extracellular signaling, APP cytoplasmic signaling). In a preferred embodiment, this screening is accomplished without the concomitant identification of inhibitors of α -secretase. The assay system is carried out in a single type of cell, using a single type of assay readout. Inhibitors discovered by means of the present invention are expected to be useful in the treatment of Alzheimer's disease since these inhibitors are likely to be capable of interfering with the production of A β .

Previous assays for identifying inhibitors of APP processing have focussed specifically on inhibition of either β -secretase or γ -secretase activity, or on inhibition of some other single aspect of A β production. In contrast, the assays described herein are directed to inhibition of APP processing in general. Substances identified through these assays may target β -secretase, γ -secretase, modulators of β -secretase or γ -secretase activity, or even an as-yet-undiscovered ligand interaction with APP. In certain embodiments, these assays will also be free of the potentially misleading or obscuring effects of α -secretase activity. In addition, unlike other assays currently in use, these assays are homogeneous assays; *i.e.*, they require no cumbersome or time-consuming steps such as column chromatography separations, immunoprecipitations, washing steps, etc. Therefore, the assays are very well adapted to a high throughput screening format.

In the present invention, novel recombinant DNA molecules are constructed in which nucleotide sequences encoding at least a portion of the luminal (i.e., N-terminal to the transmembrane region) and transmembrane regions of APP are fused to nucleotide sequences encoding a transcription factor. In a preferred 5 embodiment, the APP contains an α -secretase cleavage site that has been altered to reduce or eliminate α -secretase cleavage. This allows the assays of the present invention to avoid identifying inhibitors of α -secretase and permits the more efficient detection of β -secretase inhibitors since α -secretase and β -secretase compete for APP cleavage. The recombinant DNA molecules may be transfected, along with a reporter 10 gene, into a cell line that processes APP into A β , and stable clones may be generated. Alternatively, the recombinant DNA molecules and reporter plasmid may be utilized in transient transfections.

Upon expression in cells, the APP/transcription factor fusion protein localizes to a non-nuclear membrane of the cell (e.g., the endoplasmic reticulum) due 15 to the presence of the APP sequences in the fusion protein. In a manner similar to cleavage of APP, the fusion protein will then be cleaved, first by β -secretase and then by γ -secretase. γ -secretase cleavage releases the transcription factor from the membrane in which the APP/transcription factor fusion protein had been embedded, after which the transcription factor translocates to the nucleus and stimulates 20 transcription of the reporter gene. Assuming no α -secretase cleavage, cleavage by both β -secretase and γ -secretase is required for release of the transcription factor and transactivation of the reporter gene in this assay since γ -secretase cleavage of APP is dependent on a short luminal domain, such as that generated by β - or α -secretase cleavage. Detection of a signal from the reporter gene product will thus serve as 25 evidence of APP processing. In particular, since activation of the reporter gene requires both β -secretase and γ -secretase cleavage, the assay is capable of identifying inhibitors of both or either of these proteases.

Figure 27 is a schematic diagram depicting general features of the assay. The vertical bar in Figure 27A represents the fusion protein; the horizontal bar 30 represents the non-nuclear membrane in which the fusion protein is embedded before processing. Figure 27B shows how the transcription factor portion of the fusion protein (with small amounts of the APP portion flanking it) has moved to the nucleus following release from the fusion protein by APP processing. In the nucleus, the

transcription factor is shown binding to a regulatory DNA sequence ("Transcription Factor Response Element") and activating transcription of the reporter gene.

The recombinant DNA molecules encoding the APP/transcription factor fusion protein and the reporter gene can be used to develop novel homogenous 5 cell-based assays for the identification and assessment of inhibitors of APP processing which will be readily amenable to high throughput technology.

In one embodiment, the recombinant DNA molecules used in this invention comprise sequences encoding the amino terminal 651 amino acids of the 10 695 amino acid version of APP (Kang et al., 1987, Nature 325:733-736), including all the sequences necessary for the production of A β , as well as the C-terminal 32 amino acids of APP. The transcription factor is placed between the N-terminal and C-terminal portions of APP. The APP sequence may include a modification to increase the amount of β -secretase cleavage of the fusion protein. This modification involves mutating the K at position 612 of the α -secretase cleavage site to a V (K612V). Since 15 α -secretase and β -secretase compete for APP cleavage, reducing or eliminating APP cleavage by α -secretase results in increased β -secretase cleavage, and allows the assay to detect β -secretase inhibitors more readily. In addition, the β -secretase cleavage site within APP (KM↓DA) (SEQ ID NO:34) may be modified, e.g., to that of a naturally occurring mutation (termed the "Swedish" mutation or NL↓DA) (SEQ ID NO:38) 20 which has been shown to enhance β -secretase cleavage six-fold in cultured cells. Another possible modification is to replace the (KM↓DA) (SEQ ID NO:34) wild-type β -secretase cleavage site with the sequence (NF↓EV) (SEQ ID NO:40). The presence of NFEV in an amino acid sequence has been shown to enhance β -secretase cleavage by an even larger amount than the Swedish sequence. See U.S. Provisional Patent 25 Application Serial No. 60/292,591 and U.S. Provisional Patent Application Serial No. 60/316,115, the disclosures of which are incorporated herein, in their entirety.

In a preferred embodiment, HIV-1 TAT exon I has been fused between sequences encoding the first 651 amino acids of APP695 and the last 32 amino acids of APP695 (APP-TAT-APPct32). Co-transfection of an expression vector comprising 30 this construct with a reporter gene plasmid containing an HIV-1 LTR promoter that controls the transcription of a reporter gene leads to enhanced expression of the reporter gene. Other transcription factors that could be fused to APP1-651 include Gal4-VP16, the entire Gal4 protein, BIV TAT, HIV-2 TAT, SIV TAT, LexA-VP16, EBV Zta, Papillomavirus E2, or tissue or species specific homodimeric bHLH

transcription factors capable of activating transcription through specific DNA response elements, such as E12, E47, or Twist. The use of GAL4, BIV, HIV-2, or SIV TAT may be useful if it is desired to reduce the potency of the transactivator, thus reducing any background transactivation caused by non-specific cleavage of the fusion protein. To further reduce the potential for transactivation by TAT in the absence of β -secretase and γ -secretase cleavage, the TAT portion of the fusion protein may be altered to remove the N-terminal methionine and thus eliminate the possibility of aberrant translation of TAT through any potential internal ribosomal entry sites.

In some circumstances, high level expression of TAT has been found to be toxic to cells. Thus, when TAT is the transcription factor fused to APP in the methods of the present invention, it may be advantageous to utilize transient transfection with low amounts of the expression vector encoding the APP/TAT fusion protein. A set of preliminary experiments in which various amounts of the vector are transfected, in order to titrate acceptable levels of TAT, is recommended.

The reporter gene used will depend in large part upon the transcription factor fused to APP. The promoter used to drive the reporter gene will be LTR for TAT-based APP fusion proteins, or UAS (6x) for GAL4-VP16-based APP fusion proteins. In a particular embodiment, an LTR driving EGFP (enhanced green fluorescent protein, a brighter variant of GFP made by Aurora Biosciences, San Diego, CA) has been used to observe processing of an APP/TAT fusion protein. Under certain conditions, it may be desirable to use a less stable reporter, such as dsEGFP (a destabilized variant of EGFP made by Aurora Biosciences, San Diego, CA and marketed by Clontech, Palo Alto, CA) or a more potent reporter, such as β -lactamase. Alternatively, a stable HeLa cell line expressing LTR- β -galactosidase can be used. If the exquisite sensitivity of β -lactamase makes it less than optimal for a particular purpose, the LTR- β -galactosidase cell line may be exploited for this assay. Finally, under some circumstances Gal4-VP16 may prove to be optimal relative to TAT to reduce any inherent background problems associated with using the weakly but constitutively active LTR in the reporter plasmid, in which case the reporter plasmid could be UAS(6x)- β -lactamase (Aurora Biosciences, San Diego, CA).

A variety of cells are suitable for use in the methods of the present invention. Particularly preferred are eukaryotic, especially mammalian, cell lines. In particular embodiments, the cells are selected from the group consisting of: L cells L-M(TK⁻) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), HEK293 (ATCC CRL

1573), HEK293T, Raji (ATCC CCL 86), CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C127I (ATCC CRL 1616), BSC-1 (ATCC CCL 26), T24 (ATCC HTB-4), PC12 cells, Jurkat cells, H4 cells (ATCC HTB-148), and MRC-5 (ATCC CCL 171).

To make the assay more amenable for ultra-high throughput screening, a non-adherent cell line, such as Jurkat, can be used.

Generally, the assays of the present invention employ cells that naturally express β -secretase and γ -secretase. However, it is possible to practice the invention in cells that lack the expression of one, or both, of these enzymes. In such cases, β -secretase and γ -secretase activity can be provided by the recombinant expression of these enzymes in the cells.

In one embodiment, the present invention provides a recombinant cell, preferably a eukaryotic cell, even more preferably a mammalian cell, and most preferably a human cell, where the cell expresses a fusion protein of APP and a transcription factor and the cell contains a reporter gene that can be activated by the transcription factor. The fusion protein comprises a portion of APP where that portion includes the regions of the β -secretase and γ -secretase cleavage sites fused to a transcription factor. The region of APP including the β -secretase and γ -secretase cleavage sites can be, e.g., a portion of APP that includes amino acids 589-651 of the 695 amino acid version of APP. This region is shown below.

25 EEISEVKM DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV IA
TVIVITLVMLKKK (SEQ ID NO:33)

The β -secretase cleavage site is shown at position 596-597 (KM DA) (SEQ ID NO:34).

Two predominant cleavage sites of γ -secretase are shown at positions 636-637 and
30 638-639

(GVV IA TV) (SEQ ID NO>35).

The fusion protein will be anchored in the membrane by the APP sequences shown above. The N-terminal portion of APP must include at least the β -

secretase cleavage site, and possibly several amino-acids N-terminal to the β -secretase cleavage site to make the assay sensitive to both β -secretase and γ -secretase inhibitors. In many cases, the APP sequences will include sequences further N-terminal than those shown above, including the signal sequence at the N-terminus of APP. In cases, 5 where the APP signal sequence is not used, another signal sequence may be incorporated in the fusion protein. Such other signal sequences are known in the art.

In a related embodiment, the present invention provides a recombinant cell, preferably a eukaryotic cell, even more preferably a mammalian cell, and most preferably a human cell, where the above-described APP portion of the fusion protein contains a K612V mutation. The APP portion of this embodiment is shown below.

↓ ↓
EEISEVKM DAEFRHDSGYEVHHQVLVFFAEDVGSNKGAIIGLMVGGVV IA
TVIVITLVMLKKK (SEO ID NO:36)

15

The β -secretase cleavage site is shown at position 596-597 (KM DA) (SEQ ID NO:34).

Two predominant cleavage sites of γ -secretase are shown at positions 636-637 and 638-639

20

(GVV JA TV) (SEQ ID NO:35).

The underlined V at position 612 shows the change in sequence in the present invention from the wild-type K to the mutant V, which change provides for reduced cleavage by α -secretase.

25

In a related embodiment, the present invention provides a recombinant cell, preferably a eukaryotic cell, even more preferably a mammalian cell, and most preferably a human cell, where the above-described APP portion of the fusion protein contains the Swedish version of the β -secretase cleavage site as well as a K612V mutation. The APP portion of this embodiment is shown below.

30

↓ ↓
EEISEVNL DAEFRHDSGYEVHHQVLVFFAEDVGSNKGAIIGLMVGGVV IA
TVIVITLVMLKKK (SEQ ID NO:37)

35

35 The β -secretase cleavage site is shown at position 596-597 (NL DA) (SEQ ID NO:38).

Two predominant cleavage sites of γ -secretase are shown at positions 636-637 and 638-639

↓

(GVV IA TV) (SEQ ID NO:35).

- 5 The underlined V at position 612 shows the change in sequence in the present invention from the wild-type K to the mutant V, which change provides for reduced cleavage by α -secretase.

In a related embodiment, the present invention provides a recombinant cell, preferably a eukaryotic cell, even more preferably a mammalian cell, and most preferably a human cell, where the above-described APP portion of the fusion protein contains the NFEV version of the β -secretase cleavage site as well as a K612V mutation. The APP portion of this embodiment is shown below.

↓

↓ ↓

15 EEISEVNP EVEFRHDSGYEVHHQVLVFFAEDVGSNKGAIIGLMVGGVV IA
TVIVITLVMKKK (SEQ ID NO:39)

↓

The β -secretase cleavage site is shown at position 596-597 (NF EV) (SEQ ID NO:40).

- 20 Two predominant cleavage sites of γ -secretase are shown at positions 636-637 and
638-639

↓ ↓

(GVV IA TV) (SEQ ID NO:35).

- The underlined V at position 612 shows the change in sequence in the present invention from the wild-type K to the mutant V, which change provides for reduced cleavage by α -secretase.

The presence of both β -secretase and γ -secretase cleavage sites in the fusion proteins permits the assays of the present invention to detect inhibitors of both β -secretase and γ -secretase.

- 30 The recombinant host cells of the present invention can be further engineered to comprise a reporter gene construct. The reporter gene construct contains a reporter gene in operable linkage with a regulatory DNA sequence that confers on the reporter gene the property of being regulated by the transcription factor of the fusion protein. This regulation is such that expression of the reporter gene is
35 low or absent without binding of the transcription factor to the regulatory DNA

sequence but, when the transcription factor is released from the fusion protein by APP processing, the transcription factor can move into the nucleus of the cell and bind to the regulatory DNA sequence, thereby activating transcription from the reporter gene.

Reporter genes desirably give rise to gene products which can be
5 detected or quantitated, either in terms of amount of protein synthesized, enzymatic activity, fluorescence, luminescence, or some other phenotype. Suitable reporter gene products include firefly luciferase (de Wet et al., 1987, Mol. Cell. Biol. 7:725-737) or bacterial luciferase (Englebrecht et al., 1985, Science 227:1345-1347; Baldwin et al., 1984, Biochem. 23:3663-3667), β -lactamase, β -glucuronidase, β -galactosidase, green
10 fluorescent proteins, enhanced green fluorescent protein, destabilized enhanced green fluorescent protein, red fluorescent protein, yellow fluorescent protein, cyan fluorescent protein, destabilized yellow fluorescent protein, destabilized cyan fluorescent protein, aequorin, chloramphenicol acetyl transferase (Alton & Vapnek, 1979, Nature 282:864-869), rat liver alkaline phosphatase (Toh et al., 1989, Eur. J.
15 Biochem. 182:231-237), human placental secreted alkaline phosphatase (Cullen & Mallim, 1992, Meth. Enzymol. 216:362-368), and horseradish peroxidase, among others.

A preferred reporter gene is green fluorescent protein (GFP) or a modified GFP. Wild-type GFP has long been used in the art. Starting from green
20 fluorescent protein, many modified versions have been derived with altered or enhanced spectral properties as compared with wild-type GFP. See, e.g., U.S. Patent No. 5,625,048; International Patent Publication WO 97/28261; International Patent Publication WO 96/23810. Useful are the modified GFPs W1B and TOPAZ, available commercially from Aurora Biosciences Corp., San Diego, CA. W1B
25 contains the following changes from the wild-type GFP sequence: F64L, S65T, Y66W, N146I, M153T, and V163A (see Table 1, page 519, of Tsien, 1998, Ann. Rev. Biochem. 67:509-544). TOPAZ contains the following changes from the wild-type GFP sequence: S65G, V68L, S72A, and T203Y (see Table 1, page 519, of Tsien, 1998, Ann. Rev. Biochem. 67:509-544). Wild-type nucleotide and amino acid
30 sequences of GFP are shown in Figure 1 and SEQ ID NO: 1 of International Patent Publication WO 97/28261; in Figure 1 of Tsien, 1998, Ann. Rev. Biochem. 67:509-544; and in Prasher et al., 1992, Gene 111:229-233.

When expressing GFPs in mammalian cells, it may be advantageous to construct versions of the GFPs having altered codons that conform to those codons

preferred by mammalian cells (Zolotukhin et al., J. Virol. 1996, 70:4646-46754; Yang et al., 1996, Nucl. Acids Res. 24:4592-4593). Another way of improving GFP expression in mammalian cells is to provide an optimal ribosome binding site by the use of an additional codon immediately after the starting methionine (Crameri et al., 5 1996, Nature Biotechnology 14:315-319).

- Transcription factors that are useful in the present invention are preferably those transcription factors that are not naturally expressed in the recombinant host cells. This is so the regulatory DNA sequence is not activated absent APP processing and release of the transcription factor from the fusion protein.
- 10 Preferably, the transcription factor contains, or is engineered to contain, a nuclear localization signal. This is so that, after release, the transcription factor will move into the nucleus of the genetically modified host cells where it can bind to, and activate, the regulatory DNA sequence, leading to expression of the reporter gene. Transcription factors, as used in the present invention, do not include proteins that, 15 after release from a fusion protein and translocation into the nucleus, repress transcription from a reporter gene.

Among the transcription factors that are useful in the present invention are: HIV1 TAT (in particular exon I of HIV1 TAT), Gal4-VP16, the entire Gal4 protein, BIV TAT, HIV-2 TAT, SIV TAT, LexA-VP16, EBV Zta, Papillomavirus E2, 20 or one of the bHLH homodimeric transcription factors, E12, E47, or Twist.

Expression vectors are generally used to express the fusion protein in the recombinant cells. An expression vector contains recombinant nucleic acid encoding a polypeptide (e.g., an APP/transcription factor fusion protein) along with regulatory elements for proper transcription and processing. Generally, the regulatory 25 elements that are present in an expression vector include a transcriptional promoter, a ribosome binding site, a transcriptional terminator, and a polyadenylation signal. Other elements may include an origin of replication for autonomous replication in a host cell, a selectable marker, a limited number of useful restriction enzyme sites, and a potential for high copy number.

30 A variety of expression vectors are known in the art and can be used in the present invention. Commercially available expression vectors which are suitable include, but are not limited to, pMC1neo (Stratagene), pSG5 (Stratagene), pcDNAI and pcDNAIamp, pcDNA3, pcDNA3.1, pCR3.1 (Invitrogen, San Diego, CA), EBO-pSV2-neo (ATCC 37593), pBPV-1(8-2) (ATCC 37110), pdBPV-MMTneo(342-12)

(ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), pCI.neo (Promega), pTRE (Clontech, Palo Alto, CA), pV1Jneo, pIRESneo (Clontech, Palo Alto, CA), pCEP4 (Invitrogen, San Diego, CA), pSC11, and pSV2-dhfr (ATCC 37146). The choice of vector will depend upon the cell type in which it is desired to express the APP/transcription factor fusion protein, as well as on the level of expression desired, and the like.

The expression vectors can be used to transiently express or stably express the fusion protein. The transient expression or stable expression of transfected DNA is well known in the art. See, e.g., Ausubel et al., 1995, 10 "Introduction of DNA into mammalian cells," in Current Protocols in Molecular Biology, sections 9.5.1-9.5.6 (John Wiley & Sons, Inc.).

The recombinant host cells of the present invention are useful in methods of screening substances for the ability to inhibit APP processing. In one embodiment, the methods of the present invention comprise adding a candidate 15 substance to a recombinant host cell comprising an APP/transcription factor fusion protein and a reporter gene and comparing the level of expression of the reporter gene protein in the presence and absence of the candidate substance, wherein the level of expression of the reporter gene protein is lower when the candidate substance inhibits processing of the APP/transcription factor fusion protein such that the transcription 20 factor is not released, or is released in a lower amount, than in the absence of the substance.

The level of expression of the reporter gene protein is generally not measured directly. Rather, an indirect method is used. For example, fluorescence given off by the reporter gene protein may be detected or measured as, e.g., when the 25 reporter gene product is a green fluorescent protein; or, some enzymatic activity of the reporter gene product may be detected or measured, e.g., when the reporter gene product is β -lactamase.

The candidate substance may be of any form suitable for entry into the cytoplasm of the recombinant cell or for contact with the cell's cytoplasmic 30 membrane. Under appropriate conditions, the candidate substance may be allowed to freely diffuse into the cell, or the delivery of the substance may be facilitated by techniques and substances which enhance cell permeability, a wide variety of which are known in the art. Methods for increasing cell permeability include, without limitation, the use of organic solvents such as dimethylsulfoxide, liposomes,

application of electrical current, and physical means such as substance-coated teflon pellets.

The present invention provides a method of identifying a substance that inhibits APP processing comprising:

- 5 (a) providing a recombinant eukaryotic cell which:
 - (i) expresses a fusion protein comprising amino acids 589-651 of APP695 and a transcription factor where the transcription factor is fused in frame to the carboxyl terminus of amino acids 589-651 of APP695; and
 - (ii) comprises a reporter gene operably linked to a regulatory DNA sequence which is capable of being activated by the transcription factor;
- 10 (b) measuring the level of reporter gene product in the cell in the absence of the substance;
- 15 (c) adding the substance to the cell and measuring the level of reporter gene product in the cell in the presence of the substance;
where a decrease in the level of reporter gene product in the presence as compared to the absence of the substance indicates that the substance inhibits APP processing.

The manner in which the level of the reporter gene product is measured will be determined by the nature of the reporter gene and, often, the characteristics of 20 the host cell. For example, if the reporter gene product itself is fluorescent, as for example, when a green fluorescent protein is the reporter gene product, fluorescence from the cell can be measured directly. When the reporter gene product has enzymatic activity, for example, when the reporter gene product is β -lactamase, known methods of measuring that enzymatic activity can be used.

25 For the sake of clarity, the above method is described in terms of "a" cell. In actual practice, the method will generally be carried on a large number of cells at one time. For example, the method will often be carried out in a well of a tissue culture plate, where, depending on the number of wells in the plate (and thus their size), there can be up to hundreds, thousands, or even several million cells. The step 30 of "adding the substance to the cell" is generally carried out by simply adding the substance to the tissue culture medium in which the cells are present. After the substance is added to the cell, the cell and the substance are incubated for a period of time sufficient for the substance to inhibit APP processing, if the substance is actually

an inhibitor of APP processing. This period is usually from about 15 minutes to 48 hours, but may be somewhat more in unusual cases.

A convenient way of carrying out the method is to grow a population of the recombinant eukaryotic cells and then split the population into a portion that 5 will be exposed to the substance and a portion that will not be exposed to the substance.

The recombinant eukaryotic cell is generally produced by transfection of an expression vector encoding the fusion protein and by transfection of a plasmid containing the reporter gene.

10 One skilled in the art would recognize that what is sought in terms of "a decrease in the level of reporter gene product in the presence as compared to the absence of the substance" is a non-trivial decrease. For example, if in the method described above there is found a 1% decrease, this would not indicate that the substance is an inhibitor of APP processing. Rather, one skilled in the art would 15 attribute such a small decrease to normal experimental variance. What is looked for is a significant decrease. For the purposes of this invention, a significant decrease fulfills the usual requirements for a statistically valid measurement of a biological signal. For example, depending upon the details of the embodiment of the invention, a significant decrease might be a decrease of at least 10%, preferably at least 20%, 20 more preferably at least 50%, and most preferably at least 90%.

In particular embodiments, amino acids 589-651 of APP695 contain a K612V mutation.

In particular embodiments, the cell is a mammalian cell. In particular embodiments, the cell is a human cell.

25 In particular embodiments, the method is used to screen a library of more than 1,000 substances. In other embodiments, the method is used to screen a library of more than 50,000 substances at a rate of more than 1,000 substances per 24 hours.

30 In particular embodiments, the fusion protein comprises a portion of APP that is selected from the group consisting of: amino acids 1-651 of APP695, amino acids 50-651 of APP695, amino acids 100-651 of APP695, amino acids 150-651 of APP695, amino acids 200-651 of APP695, amino acids 250-651 of APP695, amino acids 300-651 of APP695, amino acids 350-651 of APP695, amino acids 400-

651 of APP695, amino acids 450-651 of APP695, amino acids 500-651 of APP695, and amino acids 550-651 of APP695.

In related embodiments, the fusion protein does not comprise all of amino acids 589-651 of APP695. Rather, the fusion protein comprises slightly fewer 5 amino acids from APP. For example, the fusion protein might comprise slightly fewer amino acids of the β -secretase cleavage site: e.g., amino acids 590-651 of APP695. Or the fusion protein might comprise slightly fewer amino acids of the γ -secretase cleavage site: amino acids 589-650 of APP695; amino acids 589-649 of APP695; amino acids 589-648 of APP695; or amino acids 589-647. The fusion protein may 10 even comprise slightly fewer amino acids from both ends, e.g., amino acids 590-647 of APP695. What is important is that the portion of APP included in the fusion protein contains both the β -secretase cleavage site and the γ -secretase cleavage site.

In particular embodiments, the transcription factor is selected from the group consisting of: HIV-1 TAT, Gal4-VP16, the entire Gal4 protein, LexA-VP16, 15 EBV Zta, Papillomavirus E2, one of the bHLH homodimeric transcription factors, including E12, E47, or Twist, or BIV TAT, HIV-2 TAT, or SIV TAT. A particular version of HIV-1 TAT suitable for use in the present invention is HIV-1 TAT exon I.

Fusion proteins suitable for use in the present invention can be selected from the group consisting of: APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:2); APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:4); APP(1-651)SW, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:6); APP(1-651)wt, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:8); APP(1-651)SW, TATexonI(M1L) APP (664-695) (SEQ ID NO:10); APP(1-651)wt, TATexonI(M1L) APP (664-695) (SEQ ID NO:12); APP(1-651)SW, Gal4-25 VP16(M1L) APP (664-695) (SEQ ID NO:14); APP(1-651)wt, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:16); APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:23); and APP(1-651)NFEV, K612V, GAL4-VP16(M1L) APP (664-695) (SEQ ID NO:25).

In some embodiments of the present invention, the amino acid 30 sequences contributed to the fusion protein by the transcription factor constitute the carboxy terminal amino acid sequences of the fusion protein. In other embodiments, the transcription factor has other sequences fused to its carboxy terminus, as in the examples herein where amino acids 664-695 of APP695 are fused to the carboxy terminus of the transcription factor and therefore constitute the carboxy terminal

amino acid sequences of the fusion protein. Other portions of APP (e.g., amino acids 652-695 of APP695) could be used instead of amino acids 664-695 of APP695. In fact, it should be possible to extend the carboxy terminus of the transcription factor with almost any amino acid sequences, providing such sequences do not interfere with 5 the ability of the transcription factor to move into the nucleus and activate transcription of the reporter gene once the transcription factor has been released from the fusion protein by the action of γ -secretase.

The present invention includes a method of identifying a substance that inhibits APP processing comprising:

- 10 (a) providing a recombinant eukaryotic cell which:
 - (i) expresses a fusion protein comprising an amino acid sequence from APP that is capable of being cleaved by both β -secretase and γ -secretase and a transcription factor where the transcription factor is fused in frame to the amino acid sequence from APP; and
- 15 (ii) comprises a reporter gene operably linked to a regulatory DNA sequence which capable of being activated by the transcription factor;
- (b) measuring the level of reporter gene product in the cell in the absence of the substance;
- 20 (c) adding the compound to the cell and measuring the level of reporter gene product in the cell in the presence of the substance;
 where a decrease in the level of reporter gene product in the presence as compared to the absence of the substance indicates that the substance inhibits APP processing.
- 25 In particular embodiments, the amino acid sequence from APP comprises:
 - 589-651 of APP695;
 - 590-651 of APP695;
 - 589-650 of APP695;
 - 590-650 of APP695;
- 30 589-649 of APP695;
- 590-649 of APP695;
- 589-648 of APP695;
- 590-648 of APP695;
- 589-647 of APP695; or

590-647 of APP695.

In related embodiments, the amino acid sequence from APP contains the amino acid sequence NLDA (SEQ ID NO:38) at the β -secretase cleavage site instead of the wild-type sequence KMDA (SEQ ID NO:34).

5 In related embodiments, the amino acid sequence from APP contains the amino acid sequence NFEV (SEQ ID NO:40) at the β -secretase cleavage site instead of the wild-type sequence KMDA (SEQ ID NO:34).

10 The portion of the fusion protein that is derived from APP may contain mutations that are known in the art. Of particular interest are mutations that result in an increased proportion of A β being made in the form of A β 1-42 rather than A β 1-40. Such mutations are disclosed in the following publications (numbering is from the 15 770 amino acid version of APP):

Swedish (K670N, M671L): Mullan et al., 1992, Nature Genet. 1:345-347.

Flemish (A692G): Hendriks et al., 1992, Nature Genet. 1:218-221; Cras et al., 1998,

15 Acta Neuropathol. (Berlin) 96:253-260.

Dutch (E693Q): Levy et al., 1990, Science 248:1124-1126.

Arctic (E693G): Nilsberth et al., 2001, Nature Neuroscience 4: 887-893.

Austrian (T714I): Kumar-Singh et al., 2000, Hum. Mol. Genet. 9:2589-2598.

French (V715M): Ancolio et al., 1999, Proc. Natl. Acad. Sci. (USA) 96:4119-4124.

20 Florida (T716V): Eckman et al., 1997, Hum. Mol. Genet. 6:2087-2089.

V717F: Murrell et al., 1991, Science 254:97-99.

V717G: Chartier-Harlin et al., 1991, Nature 353:844-846.

London (V717I): Goate et al., 1991, Nature 349:704-706.

L723P: Kwok et al., 2000, Ann. Neurol. 47:249-253.

25 I716F (also called I45F, referring to the position relative to the β -secretase cleavage site): This mutation in APP changes processing of A β almost exclusively to A β 1-42. Lichtenthaler et al., 1999, Proc. Natl. Acad. Sci. (USA) 96:3053-3058.

As with many proteins, it may be possible to modify many of the 30 amino acids of the fusion proteins described above and still retain substantially the same biological activity in terms of APP processing as for the original fusion protein. Thus, the present invention includes modified fusion proteins which have amino acid deletions, additions, or substitutions but that still retain substantially the same properties with respect to APP processing as the fusion proteins described herein. It is generally accepted that single amino acid substitutions do not usually alter the

biological activity of a protein (see, e.g., *Molecular Biology of the Gene*, Watson *et al.*, 1987, Fourth Ed., The Benjamin/Cummings Publishing Co., Inc., page 226; and Cunningham & Wells, 1989, *Science* 244:1081-1085). Accordingly, the present invention includes fusion proteins where one amino acid substitution has been made

5 in the fusion proteins described herein where the fusion proteins still retain substantially the same properties with respect to APP processing as the fusion proteins described herein. The present invention also includes fusion proteins where two or more amino acid substitutions have been made in the fusion proteins described herein where the fusion proteins still retain substantially the same properties with respect to

10 APP processing as the fusion proteins described herein. In particular, the present invention includes embodiments where the substitutions are conservative substitutions.

With the exception of Figure 18, the nucleotide and amino acid sequences of APP disclosed herein contain a minor difference compared to APP sequences that are usually reported in the literature. For the sequences disclosed herein with such a difference, the nucleotide at position 367 is an A rather than a G, as in most published APP sequences. This change results in a conservative substitution in the corresponding APP amino acid sequence. Thus, the amino acid sequences disclosed herein with such a difference have an I rather than a V at position 123. This 20 difference does not affect the properties of the fusion proteins for the purposes of the present invention. Therefore, fusion proteins having the APP sequence reported in the literature with an G at nucleotide position 367 and a V at amino acid position 123 and the fusion proteins disclosed herein with an A at nucleotide position 367 and an I at amino acid position 123 are to be considered equivalents for the purposes of the 25 present invention.

The Gal-VP16 sequences disclosed herein contain two changes from the usual published sequences. There is T to C change at nucleotide position 2131 that causes a S to P change at amino acid position 712; there is A to C change at nucleotide position 2301 that does not change the amino acid sequence. It is expected 30 that Gal-VP16 proteins containing the usual sequences reported in the literature will also be suitable for use in the present invention.

The methods of the present invention can be used to screen libraries of substances or other sources of substances to identify substances that are inhibitors of β -secretase or γ -secretase. Such identified inhibitory substances can serve as "leads"

for the development of pharmaceuticals that can be used to treat patients having Alzheimer's disease or in a prophylactic manner to prevent or delay the development of Alzheimer's disease. Such leads can be further developed into pharmaceuticals by, for example, subjecting the leads to sequential modifications, molecular modeling,
5 and other routine procedures employed in the pharmaceutical industry. The inhibitors of APP processing identified by the present invention may also be tested in animal models of Alzheimer's disease such as the various transgenic mouse models that are known in the art.

Although a wide variety of substances can be screened by the methods
10 of the present invention, preferred substances for screening are libraries of small molecule compounds. Small molecule compounds are preferred because they are more readily absorbed after oral administration, have fewer potential antigenic determinants, and are more likely to cross the blood/brain barrier than larger molecules such as nucleic acids or proteins.

15 Once identified by the methods of the present invention, the candidate small molecule compounds may then be produced in quantities sufficient for pharmaceutical testing and formulated in a pharmaceutically acceptable carrier (see, e.g., Remington's Pharmaceutical Sciences, Gennaro, A., ed., Mack Publishing, 1990, for suitable methods). The candidate compounds may be administered to cell lines
20 relevant to Alzheimer's disease, animal models of Alzheimer's disease, or Alzheimer's disease patients.

The numbering of the amino acids in APP used herein is based on the 695 amino acid version of APP described in Kang et al., 1987, Nature 325:733-736. There are two other major versions of APP, having 751 amino acids and 770 amino
25 acids (see, Ponte et al., 1988, Nature 331:525-527 for the 751 amino acid version and Kitaguchi et al., 1988, Nature 331:530-532 for the 770 amino acid version). One skilled in the art will understand how to translate the numbering used herein, based on the 695 amino acid version of APP, into the corresponding numbering for other versions of APP. For example, some of the APP/transcription factor fusion proteins
30 of the present invention contain the K612V mutation, based on the numbering of the 695 amino acid version. This mutation would correspond to a K668V mutation in the 751 amino acid version and a K687V mutation in the 770 amino acid version.

Therefore, when a "K612V" mutation is referred to herein, it will be understood that such reference also includes a K668V mutation of the 751 amino acid version of APP as well as a K687V mutation of the 770 amino acid version of APP.

Similarly, the portion of APP referred to as APP₁₋₆₅₁ herein, based on 5 the 695 amino acid version, will be understood to mean also APP₁₋₇₀₇ of the 751 amino acid version and APP₁₋₇₂₆ of the 770 amino acid version.

If desired, inhibitors that are identified by the methods of the present invention can be further tested to determine which step in APP processing they affect. Assays that are known to be specific for the various steps of APP processing can be 10 used for this purpose. For example, the assay of Karlström et al., (Journal of Biological Chemistry papers in press, published on December 13, 2001 as Manuscript C100649200) is only capable of detecting inhibitors of γ -secretase and cannot also detect inhibitors of other steps of APP processing such as, e.g., inhibitors of β -secretase. If an inhibitor identified by the methods of the present invention is found to 15 also be an inhibitor when tested in the assay of Karlström et al., then that inhibitor is at least a γ -secretase inhibitor. It is still possible that that inhibitor could inhibit other steps in APP processing as well. Further tests known in the art can determine this.

The present invention may be modified so as to provide methods of determining at which step of APP processing a known inhibitor of APP processing 20 exerts its effect. The known inhibitor may be one that has been identified by the methods of the present invention or by some other method. The modification to the present invention consists in mutating the β -secretase site in a fusion protein so that β -secretase cleavage can no longer occur at the site or occurs at a very much reduced level. Providing that the fusion protein contains a cleavable α -secretase site, the 25 fusion protein can still be used in the methods of the present invention. However, this fusion protein (with a mutated β -secretase site) can no longer detect β -secretase inhibitors. Therefore, if the known APP processing inhibitor still functions as an APP processing inhibitor in this modified version of the invention, then the known inhibitor cannot be a β -secretase site inhibitor but instead must exert its effect 30 downstream of β -secretase.

Suitable mutations of the β -secretase site include the following. All the sequences are for amino acid positions 594-598 of APP695.

VNFAV (SEQ ID NO:41): This mutation shows decreased β -secretase cleavage relative to the wild type, KMDA (SEQ ID NO:34), sequence.

VKVDA (SEQ ID NO:42): Vassar et al., 1999, Science 286:735-741. This mutant was tested in vitro only, but purified β -secretase failed to cleave a 30-amino acid peptide containing this sequence.

- 5 WKMDA (SEQ ID NO:43), VKADA (SEQ ID NO:44), VKKDA (SEQ ID NO:45),
VKEDA (SEQ ID NO:46), VKIDA (SEQ ID NO:47), VKMIA (SEQ ID NO:48),
VKMNA (SEQ ID NO:49), VKMEA (SEQ ID NO:50), VKMDE (SEQ ID NO:51),
VKMDK (SEQ ID NO:52): Citron et al., 1995. Neuron 14:661-670. These mutations decreased A β production 4-20X relative to p3 production in cultured cells.

Fusion proteins can be constructed by use of the polymerase chain reaction (PCR) to amplify desired portions of APP and transcription factors, which can be then be cloned into expression vectors by methods well known in the art. Primers for PCR will generally include a small part of the APP or transcription factor as well as convenient cloning sites and/or linker peptide sequences. The PCR primers can be used to amplify the desired APP or transcription factor fragments from sources such as previously cloned APP or transcription factors, cDNA libraries, or genomic libraries. The amplified APP and transcription factor sequences can be cloned into suitable expression vectors. Methods of PCR and cloning are well known in the art and can be found in standard reference materials such as those listed below.

Standard techniques for cloning, DNA isolation, amplification and purification, for enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like, and various separation techniques are known and commonly employed by those skilled in the art. A number of standard techniques are described in Sambrook et al. (1989) Molecular Cloning, Second Edition, Cold Spring Harbor Laboratory, Plainview, N.Y.; Maniatis et al. (1982) Molecular Cloning, Cold Spring Harbor Laboratory, Plainview, N. Y.; Wu (ed.) (1993) Meth. Enzymol. 218, Part I; Wu (ed.) (1979) Meth. Enzymol. 68; Wu et al. (eds.) (1983) Meth. Enzymol. 100 and 101; Grossman and Moldave (eds.) Meth. Enzymol. 65; Miller (ed.) (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.; Old and Primrose (1981) Principles of Gene Manipulation, University of California Press, Berkeley; Schleif and Wensink (1982) Practical Methods in Molecular Biology; Glover (ed.) (1985) DNA Cloning Vol. I and II, IRL Press, Oxford, UK; Hames and Higgins (eds.) (1985) Nucleic Acid Hybridization, IRL Press, Oxford, UK; Setlow and Hollaender (1979) Genetic Engineering: Principles

and Methods, Vols. 1-4, Plenum Press, New York, and Ausubel et al. (1992) Current Protocols in Molecular Biology, Greene/Wiley, New York, N.Y..

PCR reactions can be carried out with a variety of thermostable enzymes including but not limited to AmpliTaq, AmpliTaq Gold, or Vent polymerase.

- 5 For AmpliTaq, reactions can be carried out in 10 mM Tris-Cl, pH 8.3, 2.0 mM MgCl₂, 200 µM of each dNTP, 50 mM KCl, 0.2 µM of each primer, 10 ng of DNA template, 0.05 units/µl of AmpliTaq. The reactions are heated at 95°C for 3 minutes and then cycled 35 times using suitable cycling parameters, including, but not limited to, 95°C, 20 seconds, 62°C, 20 seconds, 72°C, 3 minutes. In addition to these
10 conditions, a variety of suitable PCR protocols can be found in PCR Primer, A Laboratory Manual, edited by C.W. Dieffenbach and G.S. Dveksler, 1995, Cold Spring Harbor Laboratory Press; or PCR Protocols: A Guide to Methods and Applications, Michael et al., eds., 1990, Academic Press.

- 15 It is desirable to sequence the DNA encoding the fusion proteins, or at least the junction regions of the various portions (APP, transcription factor, linkers) of the fusion protein in order to verify that the desired portions have in fact been obtained, joined properly, and that no unexpected changes have been introduced into the sequences by the PCR reactions.

- 20 Suitable PCR primers for amplification of DNA sequences for use in the present invention can be readily designed by those of skill in the art. Examples of such primers are shown below.

5'-GGA GAG GAT ATC ATG GAG CCA GTA GAT CC-3' (SEQ ID NO:53) can be used to amplify the 5' portion of HIV-1 TAT exon I.

- 25 5'-TAC ATG GCG GCC GCC TAC TTA CTG CTT TG-3' (SEQ ID NO:54) can be used to amplify the 3' portion of HIV-1 TAT exon I.

- 30 5'-GGA TGT GAT ATC TTT CTT CTT CAG CAT CAC CAA GG-3' (SEQ ID NO:55) can be used to amplify the 3' portion of DNA encoding amino acids 1-651 of APP, i.e., the transmembrane region of APP.

The following non-limiting examples are presented to better illustrate the invention.

EXAMPLE 1

Transfection of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI with
pMM321

5 The following example demonstrated that an APP/TAT fusion construct will transactivate a reporter gene in which the HIV1 LTR regulatory DNA sequence controls the expression of enhanced green fluorescent protein (EGFP). The following also serves as an example of the kind of preliminary routine variations of fusion protein levels and inhibitor levels that may be advantageous to test in the
10 practice of the present invention. Such routine variations are often helpful in validating the assays before a large scale screening project is undertaken.

15 The APP/TAT fusion construct is referred to as “pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI” (see Figure 26) and contains the HIV1 TAT exon 1 fused just after the transmembrane domain of APP. This construct is also shown in outline form in Figure 1B. “pMM321” refers to a reporter gene plasmid consisting of the HIV1 LTR driving the transcription of enhanced green fluorescent protein (see Figure 25). As a positive control for TAT expression, a construct in which TAT was under the control of a strong, constitutive promoter (referred to as “pUCd5TAT”; see Figure 24) was used.

20

METHODS:

1. Day 1: Pass HEK 293T cells into 2 x 6 well dishes at 1×10^5 cells/well.
2. Day 2: Transfect cells with 9 μ L Fugene and 0.125 μ g pMM321 and various amounts of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI.

- 25 Plate 1:
1. 5 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
 2. 2.5 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
 - 30 3. 1.25 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI

4. 0.625 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
- 5 5. 0.312 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
6. 0.156 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
- Plate 2:
 1. 0.08 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
 2. 0.04 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-
- 10 (M1L)TATexonI
 3. no pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
 4. 0.625 pUCd5TAT
 5. 0.625 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI and 1 µg pMM321
 - 15 6. 1 µg pMM321

Six hours post-transfection, green cells were only observed in plate 2, #5.

3. Day 3: The fluorescence intensity of the transfected cells was observed and recorded.
- 20
4. Day 4: The fluorescence intensity of transfected cells was observed and recorded.

RESULTS:

- 25 Co-transfection with pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI increased GFP expression in the cells.

Day 3:

- 30 • 5 µg – no green cells
- 2.5 µg – no green cells (too much DNA for these two transfections?)
- 1.25 µg - many bright and dim green cells (see photographs and figure in ancillary data)
- 0.625 µg - bright and dim green cells but fewer than at 1.25 µg
- 0.312 µg - no difference obvious between 0.625 µg and 0.312 µg

- 0.156 µg - very few green cells
- 0.08 µg - very few green cells
- 0.04 µg - very few green cells
- no pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI - extremely few
5 (if any) green cells
- 0.625 µg pUCd5TAT - cells were extremely bright, not necessarily more in number than with pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
- 0.625 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 1 µg
pMM321 - many bright and dim green cells.
- 10 • 1 µg pMM321 many-fold fewer green cells, some bright, most dim.

Day 3 - changed media (saved 1 mL conditioned media from wells Plate 1- 3, 4, 5, 6; Plate 2 - 1, 2, 3, 5, 6). Added fresh media with 10 µM of L-685,458 (a potent, cell permeable γ- secretase inhibitor) to wells Plate 1 - 3, 4, 5, 6; Plate 2 - 3, 4, 5, 6.
15 Waited 48 hours to observe loss of fluorescence since GFP is so stable.

After 48 hours, all wells appeared brighter than at 24 hour time point. This does not necessarily mean that the inhibitor was ineffective, or that the assay did not work, since there were no controls run where the inhibitor was not added. However, it does
20 suggest that under these conditions it may be preferable to add the inhibitor at the time of transfection to shut down γ-secretase as soon as possible and avoid release of TAT and induction of GFP.

EXAMPLE 2

25 Transfection of APP(1-651)SW, K612V-(M1L)TATexonI into HEK293T and H4
cells accompanied by inhibition of γ-secretase activity with L-685,458

The following example demonstrates the operation of the invention in HEK293T cells and H4 cells and shows inhibition of APP processing (and thus TAT release) by treatment with a known γ-secretase inhibitor. "pcDNA3.1 zeo (+) APP(1-
30 651)SW, K612V-(M1L)TATexonI," "pMM321," and "pUCd5TAT" are the same as in Example 1. H4 cells (ATCC HTB-148) are a neuronal cell line.

METHODS:

Day1: Plated out 2 x 6 well plates of HEK293T cells and 2 x 6 well plates of H4 cells at 1×10^5 cells/well.

5 Day 2: Transfected cells with 2 μ g total DNA - pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI and carrier (a pET-IN plasmid).

Plate1:

1,2: 1 μ g pMM321 + 1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI

10 3,4: 1 μ g pMM321 + 0.1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 0.9 μ g carrier

5: 1 μ g pMM321 + 1 μ g carrier (added too much carrier to this well in H4 cells)

6: 1 μ g pMM321 + 0.1 μ g pUCd5TAT + 0.9 μ g carrier

15 Plate 2:

1,2: 0.1 μ g pMM321 + 1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 0.9 μ g carrier

3,4: 0.1 μ g pMM321 + 0.1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI +

20 1.8 μ g carrier (added too 0.5X carrier to this mix in 293T cells)

5: 0.1 μ g pMM321 + 1.9 μ g carrier

6. 0.1 μ g pMM321 + 0.1 μ g pUCd5TAT + 1.8 μ g carrier

Transfections for HEK293T cells: 9 μ L Fugene/well. Combined with DNA in Optimem and incubated and added to cells according to manufacturer's instructions.

Transfections for H4 cells: 6 μ L Fugene/well. Combined with DNA in Optimem and incubated and added to cells according to manufacturer's instructions.

30 Added 10 μ M L-685,458 to Plates 1 and 2, wells 2 and 4 for both cell types within 1 hour of transfection. Observed cells periodically.

Took pictures at 24, 46 hours after transfection, using AE lock to keep exposures constant between wells.

RESULTS:

Both H4 and 293T cells turned much brighter green in the presence of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI

5 At 24 hours:

H4 cells:

Plate 1: 1 ug pMM321

1. 1 ug pMM321 + 1 ug pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: Many

10 bright and dim green cells (good transfection efficiency as well)

2. 1 ug pMM321 + 1 ug pcDNA3.1 zeo (+) APP(1-651)SW, K612V-

(M1L)TATexonI + 10 μ M L-685,458: Also many bright and dim green cells, but reduced compared with

well #1

15 3. Very few green cells (a few per field)

4. Very few green cells

5. A few dimly green cells

6. Some induction with 0.1 μ g pUCd5TAT but still relatively few cells.

20 Plate 2: 0.1 μ g pMM321

1. 0.1 μ g pMM321 + 1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI:

~10 bright green cells/field and the rest are dim green

2. 0.1 μ g pMM321 + 1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-

25 (M1L)TATexonI + 10 μ M L-685,458: ~3-5 bright green cells/field, some dim green, and some not green.

3. No green cells

4. No green cells

5. No green cells

30 6. A few bright green cells

HEK293T cells:

Plate 1: 1 μ g pMM321

1. + 1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: of 15 cells:
6 dim, 4
medium, 5 bright
2. + 1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 µM L-
5 685,458:
of 15 cells: 6 very dim, 5 dim, 1 medium, 3 bright
3. +0.1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: many more
green
cells than 1 µg, lots of strong, bright green cells
- 10 4. + 0.1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 µM L-
685,458:
Fewer bright green cells/field but intensity does not appear strongly diminished
5. 1 µg pMM321 alone: Most cells in the field expressing dim to medium levels of
GFP
- 15 6. enhancement by 0.1 µg pUCd5TAT

- Plate 2: 0.1 µg pMM321
1. + 1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: Bright and
medium
20 green cells
 2. + 1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 µM L-
685,458:
Bright, medium, and dim green cells
 3. + 0.1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI: Bright,
25 medium,
and dim green cells
 4. + 0.1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 10 µM L-
685,458:
Bright, medium, and dim green cells
 - 30 5. 0.1 µg pMM321 alone: Most expressing cells have dim GFP, a few medium to
bright
cells
 6. Enhancement by 0.1 µg pUCd5TAT

Changed media on cells at 24 hours past transfection. Kept 10 μ M L-685,458 on cells in wells 2 and 4.

At 46 hours after transfection, examined the wells again. Lots of floating cells in all 5 wells, all cell types. Highest number of floaters in 1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI lanes.

Took some photographs under fluorescent and white light (white light at low 10 intensity) to reveal fluorescent and non-fluorescent cells. Conducted a subjective analysis of the photographs to see if the amount of inhibition by 10 μ M L-685,458 was in any way quantifiable. Counted bright (white in the middle); strong (blue middle), medium (green) and dim/non-fluorescent cells and determined the approximate fraction of each level of expression. Results follow:

15

TABLE 1

293T cells transfected with X μ g pMM321 (first number in left column) and X μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI (second number in left column)

Transfection	# Bright	# Strong	# Med	# Non	% Bright	%Strong	% Med	%Non
1 μ g + 1 μ g	none	8	81	316		2	20	78
1 + 1 + cmpd	none	5	66	521		0.8	11	88
1 + 0.1	41	77	143	61	12	24	44	19
1+ 0.1+cmpd	23	32	149	194				

20 The results shown in Table 1 indicate that the presence of L-685,458 ("cmpd") caused fewer strong and medium fluorescing cells as well as more non-fluorescent cells in the first run; in the second run, L-685,458 caused fewer bright and strong fluorescing cells as well as more non-fluorescent cells (with slightly more medium fluorescing cells). Overall, these data clearly indicate that the presence of an inhibitor of APP processing 25 such as L-685,458 can be identified by the present invention.

1 mL of conditioned media from each well was analyzed for production of A β . Higher than background levels of A β were observed in 293T cells after transfection

- with 1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI and 0.1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI and higher than background levels of A β in H4 cells after transfection with 1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI, but not 0.1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI (no enhancement of GFP was observed with 0.1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI either). A β was completely inhibited to background levels by 10 µM L-685,458. Surprisingly, substantial inhibition of GFP was not observed with 10 µM L-685,458.
- 5 100,000 cells from each well were trypsinized and placed in 0.1 mL phenol red-free media in a Costar 96-well dish and read using the fluorometer. The results are shown below:

TABLE 2

A		1ugAPPTAT	10uM 458	0.1ugAPPtat	10 uM458	no tat	0.1ug pUCd5STAT						
B	293T	1ug <u>LTRGFP</u>	17670	14321	65535	65535	14890	65535					
C		0.1ug <u>LTRGFP</u>	9976	10491	17677	14790	9624	25735					
D	H4	1ug <u>LTRGFP</u>	21307	25307	7175	7136	7147	7277					
E		0.1ug <u>LTRGFP</u>	9574	10031	7317	6957	6946	7247					
F		<u>Blank</u>	7498	7124	7570	7454	5774	7638					
G													
H													
		1	2	3	4	5	6	7	8	9	10	11	12

- 15 In Table 2, "APPtat" is pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI. "458" is L-685,458. "LTRGFP" is pMM321.

0.1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI with and without compound exceeded the maximum reading of the fluorometer, as did the addition of 0.1 µg pUCd5TAT to cells transfected with 1 µg pMM321.

1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 1 µg pMM321

- 5 incrementally increased the amount of fluorescence relative to 1 µg pMM321 alone, and this was reduced to background levels by 10 µM L-685,458. Inhibition of fluorescence was also observed in 293T cells transfected with 0.1 µg pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI + 0.1 µg pMM321. No inhibition of fluorescence was observed in H4 cells under any transfection conditions.

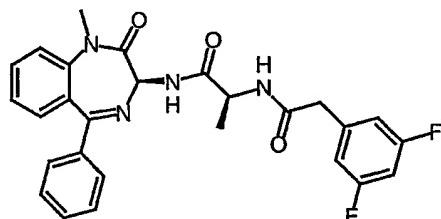
10

EXAMPLE 3

Use of APP(1-651)SW, K612V-TATexonI in H4 cells

L-875,532 is a known γ-secretase inhibitor having the structure shown below. It is described and details of its synthesis are disclosed in Seiffert et al., 2000, J. Biol.

- 15 Chem. 275:34086-34091.



L-875532

Compound X is a β-secretase inhibitor.

20

pRBR186 (Figure 22A) is an expression vector containing DNA sequences encoding full-length APP containing the Swedish mutation and the K612V mutation. pRBR186 does not contain a transcription factor fused to the APP sequences.

pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI is an expression vector that directs the expression of a the fusion protein APP(1-651)SW, K612V-(M1L)TATexonI in mammalian cells. This fusion protein contains the first 651 amino acids of APP (with a Swedish version of the β -secretase cleavage site as well 5 as the K612V mutation) fused in frame to exon I of HIV1 TAT, which has been modified with a Met1-Leu mutation. A schematic diagram of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI is shown in Figure 26A. The nucleotide sequence of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI (SEQ ID NO:22) is shown in Figure 26B-G.

10

METHODS:

1. H4 cells (ATCC HTB-148) were transfected with the various constructs listed below using 6 μ L Fugene per 100 μ L Optimem and 100 μ L Optimem per well (6-well dishes). Transfection reactions were incubated for 30 minutes prior to adding 100 μ L 15 dropwise onto wells.

Transfections were done as follows:

1. 1 μ g pMM321 (Figure 25A-D) and 1 μ g pcDNA3.1 backbone
- 20 1a. 1 μ g pMM321 and 1 μ g pcDNA3.1 (Invitrogen, San Diego, CA) backbone. Prior to transfection, 10 μ M L-875,532 (γ -secretase inhibitor) was added to the well.
2. 1 μ g pMM321 and 1 μ g pRBR186 (Figure 22A; APP expression vector; processing
- 25 and inhibition of processing control)
- 2a. 1 μ g pMM321 and 1 μ g pRBR186. Prior to transfection, 10 μ M L-875,532 was added to cells (transfection solution for 3-5 were prepared in bulk)
3. and 3a. 1 μ g pMM321 and 1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI
- 30 4. and 4a. 1 μ g pMM321 and 1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI. Prior to transfection, 10 μ M L-875,532 was added to the two wells.
5. and 5a. 1 μ g pMM321 and 1 μ g pcDNA3.1 zeo (+) APP(1-651)SW, K612V-

(M1L)TATexonI. Prior to transfection, 10 μ M Compound X was added to the two wells

6. 1 μ g pMM321 and 1 μ g pUCd5STAT (Figure 24).

7. 1 μ g pMM321 and 1 μ g pUCd5 TAT. Prior to transfection, 10 μ M L-875,532 was
5 added to the cells.

RESULTS:

Cells were assessed by eye under a fluorescence microscope the morning following transfection (~20 hrs).

10 1 and 1a, 2 and 2a. Weak fluorescence

3 and 3a. Much stronger fluorescence

4 and 4a. Clear inhibition of fluorescence

5 and 5a. Possible inhibition of fluorescence, but doesn't look that great

6 and 7. Almost blindingly fluorescent.

15

At approximately 48 hours, cells were trypsinized, spun down, and resuspended in 100 μ L PBS. The cellular contents of each well of the transfection plates were placed into one well of a 96-well fluorescence plate. Fluorescence was analyzed using the FLUOstar (485 excitation/538 emission). The results are shown in Table 3.

20

TABLE 3

Transient transfections in H4 cells	Fluor Units
pMM321	4484
pMM321 + L-875,532	3443
pMM321 + pRBR186	2735
pMM321 + pRBR186 + L-875,532	2161
pMM321 + APP-TAT-ct32	20177
pMM321 + APP-TAT-ct32 + L-875,532	8283
pMM321 + APP-TAT-ct32 + Compound X	11946
pMM321 + pucd5-TAT	61102

In Table 3, "APP-TAT-ct32" refers to pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI.

For a graphical presentation of these results, see Figure 19. In Figure 19, "LTR-GFP" refers to pMM321; "APP-TAT-ct32" refers to pcDNA3.1 zeo (+) APP(1-651)SW, K612V-(M1L)TATexonI. Compare the bar labeled "LTR-GFP + APP-TAT-ct32"

- 5 with the bars labeled "LTR-GFP + APP-TAT-ct32 + L-875,532" and "LTR-GFP + APP-TAT-ct32 + Compound X." Inhibition by both the β -secretase inhibitor (Compound X) and the γ -secretase inhibitor (L-875,532) is easily identified by the present invention.

10 Conclusions:

- The data indicate that the expression of the fusion protein APP(1-651)SW, K612V-(M1L)TATexonI enhances transactivation through the LTR of pMM321 in a manner that depends on APP processing.
- APP(1-651)SW, K612V-(M1L)TATexonI expressing cells were 6X brighter than pMM321 cells alone.
- Treatment with L-875,532 decreased fluorescence 2.5X.
- Treatment with Compound X decreased fluorescence 1.7X.
- Expression of TAT via pucd5-TAT was almost blinding and was 19X above pMM321 alone, indicating that APP(1-651)SW, K612V-(M1L)TATexonI expression did not lead to levels of GFP as high as TAT alone. Despite the decreased activation shown by TAT when provided by the fusion protein, as compared with TAT driven by the AMLP (adenovirus major late promoter) in pucd5-TAT, the assay was easily able to identify both the β -secretase and the γ -secretase inhibitors.
- Control plasmids (pMM321 and pMM321 + pRBR186) were dimly fluorescent and were not inhibited by L-875,532.

A lower level of inhibition by the β -secretase inhibitor is to be expected since the K612V mutation decreases alpha-secretase activity by 95% and thus some alpha-secretase cleavage is to be expected.

EXAMPLE 4

Construction of pcDNA3.1 (+) zeo APP(1-651)SW, K612V, GAL4-VP16(M1L) APP
(664-695)

1. The GAL4-VP16 insert was prepared by PCR from pCR2.1 GAL4VP16 (Figure
5 20) (Invitrogen, San Diego, CA). The PCR was performed to eliminate the N-
terminal methionine by changing this methionine into a leucine.

40 ng pCR2.1 GAL4-VP16

0.2 µL GAL4VP16 5' oligo at 250 µM:

10 5'-CTGAGATATCAAGCTACTGTCTTCTATCGAACAAAGC-3' (SEQ ID NO:56)

EcoRV site underlined

0.2 µL GAL4VP16 3' oligo (at 250 µM): 5'-

GCGCGATATCCCCACCGTACTCGTCAATTCC-3' (SEQ ID NO:57)

EcoRV site underlined

15 5 µL 10X Buffer

8 µL 25 mM MgCl₂

4 µL PCR dNTPs

0.25 µL AmpliTaq Gold

27.35 µL water

20

Cycle:

Purified reactions using a Qiaquick column

Digested entire reaction using EcoRV

25 Ran the DNA on a 1% gel. Excised the band and purified using a QiaQuick gel
purification kit

2. Digested pcDNA3.1 APP(1-651)/APP(664-695) with EcoRV and SAP treated.
pcDNA3.1 APP(1-651)/APP(664-695) is an intermediate plasmid formed in the
30 procedure described in Example 6. pcDNA3.1 APP(1-651)/APP(664-695) the first
651 amino acids of APP (with a Swedish version of the β-secretase cleavage site as
well as the K612V mutation) fused in frame to the last 32 amino acids of APP.

3. Ligated pcDNA3.1 APP(1-651)/APP(664-695) -EcoRV digested to GAL4VP16
(EcoRV digested)

5 4. At this point, it was realized that the 3' PCR primer for GAL4-VP16 put the
APP(664-695) fragment out of frame. The APP(664-695) fragment was then re-
PCR'd using the following protocol:

1 μL pcDNA3.1 APP(1-651)-Gal4VP16-APP(664-695)

10 50 nM APP NotI 5' ct32 in frame with GAL4-VP16

50 nM APP NotI 3' ct32 (5'

(p)CTGCTGTGGCGGCCGCCTAGTTCTGCATCTGCTC (SEQ ID NO:58)

NotI site underlined

1 μL PCR dNTPs (10 mM each dNTP, Roche)

15 5 μL 10X Expand Buffer with MgCl₂

40 μL water

1 μL Expand polymerase (Roche)

The PCR fragment was run on a 4% agarose gel and gel-purified using a QiaQuick gel
20 purification column

The fragment was digested with NotI and purified using a QiaQuick PCR purification
column.

25 5. APP(1-651)-Gal4VP16-APP(664-695) was re-miniprepped. Miniprep #1 was
digested with NotI, run on a 1% gel, the upper band was then isolated and SAP-
treated.

6. APP(1-651)-Gal4VP16/NotI digested/SAP-treated was ligated to APP(664-695).

30 7. Minipreps containing inserts were sequenced to verify the orientation of the insert.

EXAMPLE 5

Construction of pcDNA3.1 zeo (+) APP(1-651)wt, K612V-TATexonI(M1L)
APP(664-695)

- This procedure replaced a fragment of APP in pcDNA3.1 zeo (+) APP(1-651)SW,
5 K612V-TATexonI(M1L) APP (664-695) that contained the Swedish mutation with a
corresponding fragment from pRBR121 containing the wild-type β -secretase cleavage
site rather than the Swedish β -secretase cleavage site.
1. pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695)
10 (Figure 21B) was digested with SnaBI and then EcoRI
17 μ L miniprep DNA
2 μ L 10X buffer
1 μ L SnaBI (NEB)
- 15 Digest was purified using Qiaquick PCR purification kit. Entire digest was then
cleaved with EcoRI for 2 hours.
2. pRBR121 (Figure 21A) was digested with SnaBI and then EcoRI
20 5 μ g pRBR121
5 μ L 10X buffer
2.5 μ L SnaBI (NEB)
q.s. 50 μ L with water
- The digest was purified using a Qiaquick PCR purification kit. The entire digest was
25 then cleaved with EcoRI for 2 hours.
3. Both digests were run out on a 1% agarose gel. From the pRBR121 lane, the 2.4
kb SnaBI-EcoRI fragment containing the wild-type β -secretase cleavage site was
isolated.
- 30

From pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) digest, BOTH the 5 kb SnaBI-EcoRI backbone fragment AND the 200 bp EcoRI-EcoRI fragment were isolated (see Figure 21B).

- 5 4. A three-part ligation using equal molar ratios of the three fragments was carried out:

The assumption was made that, since the starting plasmids were of similar sizes and the same amount was digested for each plasmid, the recovered fragments would be in approximately equal molar ratios.

10

Vector alone:

1 μL APP-TAT-ct32 SnaBI/EcoRI 5Kb fragment

7 μL water

2 μL 5X buffer

15 10 μL 2X buffer (Roche rapid ligation kit)

1 μL T4 ligase

1+1+1

1 μL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695)

20 SnaBI/EcoRI 5KB fragment

1 μL pRBR121 SnaBI/EcoRI 2.4 Kb insert

1 μL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695)

EcoRI/EcoRI 200 bp insert

5 μL water

25 2 μL 5X buffer

10 μL 2X buffer

1 μL T4 ligase

1+1+...1 (in this 3-pt ligation, the ligation of two of the fragments together was done
30 1st, then the third fragment was added)

1 μL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695)

SnaBI/EcoRI 5Kb backbone

1 μL pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695)

EcoRI/EcoRI 200 bp insert

- 5 μ L water
2 μ L 5X buffer
10 μ L 2X buffer
1 μ L T4 ligase
- 5 waited 5 minutes
then added 1 μ L pRBR121 SnaBI/EcoRI 2.4 kb insert
- 1+1+3
1 μ L pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695)
- 10 SnaBI/EcoRI 5 kb backbone
1 μ L pRBR 121 SnaBI/EcoRI 2.4 kb insert
3 μ L pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695)
EcoRI/EcoRI 200 bp insert
3 μ L water
- 15 2 μ L 5X buffer
10 μ L 2X buffer
1 μ L T4 ligase
- Transformed and plated out 200 μ L. The number of colonies in the vector + insert
20 ligations far exceeded the number of colonies in the vector alone ligation. Picked 12 colonies from 1+ 1+...1.
Picked 6 colonies from 1+3.
Miniprepped
Digested with EcoRI to ensure that small 200 bp fragment was incorporated.
- 25 RESULTS
Minipreps #10 and 15 contained 200 bp EcoRI fragment.
- Oriented with Bam HI digestion.
- 30 Sequenced with sAPPb F2 and F3 primers. Miniprep #15 contains both inserts in the correct orientation.

EXAMPLE 6

Construction of pcDNA3.1 zeo (+) APP(1-651)SW, K612V-TATexonI(M1L)
APP(664-695)

1. PCR of APP (664-695):
5 The starting material was the pRBR186 plasmid (Figure 22A).

PCR of APP (664-695)

- 1 ng pRBR186
50 nM 5' oligo (5'-TGCCCCGCGGGCCGCGCGATGCTGCCGG-3') (SEQ ID
10 NO:59) NotI site underlined
50 nM 3' oligo (5'-(p)ATGGTGTGGCGGCCGCAGACGCCGCTGTCACC-3')
(SEQ ID NO:60) NotI site underlined
1 µL Roche PCR nucleotides
5 µL 10X Expand buffer
15 40 µL water
1 µL Expand

Cycle: (94°C, 5 min) – 25X (94°C, 30 sec; 42°C, 1 min; 72°C, 2 min) – 72°C x 6 min
– 4°C hold.

- 20 The ~100 bp fragment was gel purified (4% agarose, 1X TBE gel)

The gel-purified fragment was ligated into NotI digested, Shrimp Alkaline Phosphatase-treated pcDNA3.1 zeo (+) (Invitrogen). The presence of the insert and
25 its orientation was confirmed by sequencing.

2. PCR of APP(1-651):
1 ng pRBR186
50 nM 5' oligo (5'-(p)AGCGCACAAGCTTCCCCGCGCAGGGTCGCGATGCTG-
3') (SEQ ID NO:61) HindIII site underlined, Met(1) ATG of APP in bold
30 50 nM 3' oligo (5'-(p)GGATGTAAGCTTTTCTTCAGCATCACCAAGG-3')
(SEQ ID NO:62) HindIII site is underlined

1 µL Roche PCR nucleotides
5 µL 10X Expand buffer
40 µL water
1 µL Expand

5

Cycle: (94°C, 5 min) -- 25X (94°C, 30 sec; 37°C, 1 min; 72°C 2.5 min) – 72°C x 6 min – 4°C hold

- The amplified fragment was isolated on an agarose gel. The fragment was purified from the gel using Qiaquick Gel purification columns. The fragment was digested with HindIII. The amount of the fragment was too small to subclone, so the PCR was repeated using 1 µL of the amplified fragment and carrying out 5 reactions simultaneously.
- 15 • The fragments were purified from these reactions using a QiaQuick PCR purification kit. The fragments were eluted in 30 µL and digested with HindIII for 2 hours. The digested fragments were then gel purified.
- 20 • The purified fragments were ligated to pcDNA3.1 zeo (+) APP(664-695) that had been digested with HindIII and SAP treated. This gave the intermediate plasmid pcDNA3.1 zeo (+) APP(1-651)/APP(664-695).

3. PCR of (M1L) TAT:

The starting material was NL4-3 viral plasmid (Figure 22B).

25

PCR reaction:

1 ng NL4-3 viral plasmid

50 nM TAT 5' RV Met-Leu PCR primer (5'-

TGCAGATATCCTGGAGCCAGTAGATCCTAGAC-3') (SEQ ID NO:63)

30 EcoRV site underlined, Met-Leu mutation in bold

50 nM TAT 3' RV Met-Leu PCR primer

(5'-**GCTGGATATCCTCTGCTTGATAGAGAAGC-3'**) (SEQ ID NO:64)

EcoRV site underlined

1 µL PCR dNTPs

5 µL PCR 10X buffer with MgCl₂

40 µL water

1 µL Expand polymerase

5 Cycle:

94°C for 5 min

[30 sec 94°C, 1 min 42°C, 1 min 72°C] x 25 cycles

5 min at 72C

hold at 4°C

10

- The insert was purified over QiaQuick PCR purification column
- The entire reaction was digested with 30 units EcoRV for 3 hours
- The ~200 bp insert was gel purified.
- pcDNA3.1 zeo (+) APP(1-651)/APP(664-695) was digested with EcoRV, and then SAP treated
- The Met1-Leu TAT fragment was ligated to pcDNA3.1 zeo (+) APP(1-651)/APP(664-695).

A map of the resulting plasmid is shown in Figure 22C.

20

EXAMPLE 7

Design of novel expression vector for expression of APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695)

PURPOSE: To provide a low level expression of APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695), a prokaryotic selectable marker that is NOT ampicillin (read-through of the β-lactamase gene is sometimes a problem), and a eukaryotic selectable marker that is NOT zeocin (zeo is the marker for the reporter plasmids in some embodiments).

METHODS

30 1. The dEYFP gene was removed from pd2EYFP (Clontech, Palo Alto, CA) using BamHI and NotI. The 5' overhangs were filled in using Klenow, and the plasmid was re-circularized.

pd2EYFP plasmid was digested with BamHI, NotI.

Ran reaction on 1% agarose gel. Digestion was complete. Cut out 3.4 kb band.

Purified using Qiagen Gel Extraction Kit.

5 Klenow fill-in:

~4 µg plasmid backbone

7.5 µL NEB buffer 2

33 µM each dNTP (diluted from Roche PCR dNTPs)

water to 75 µL

10 4 µL Roche Klenow fragment (4 units)

Incubated at room temperature for 15 minutes

Heat inactivated at 70°C

15 Took 1 µL fill-in reaction.

Diluted to 8 µL with water

Added 2 µL 5X DNA buffer

Added 10 µL 2X Ligation Buffer

Added 1 µL T4 DNA ligase

20

Incubated at room temperature

Transformed 2 µL ligation into Invitrogen maximum efficiency DH5alpha competent cells.

25

Plated out on Kanamycin plates. Lots of colonies.

2. The RSV promoter from pREP4 (Invitrogen) was excised using BglII and HindIII and cloned into the re-circularized plasmid.

30

Digested 5 µg pREP4 with HindIII

Purified using Qiagen PCR purification kit

Digested with BglII.

The RSV promoter fragment was gel purified and cloned into the re-circularized plasmid.

3. The resulting expression vector (pRSV Kan/Neo res; Figure 23) has the eukaryotic
5 RSV promoter 5' to the pd2EYFP polylinker, SV40 driving neo and kanamycin
prokaryotic selection, and a pUC ori for high levels of replication in bacteria.

EXAMPLE 8

10 Use of APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695) in HeLa cells with a
β-galactosidase reporter gene

The following demonstrates the practice of the present invention with the APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695) fusion protein (SEQ ID NO:2) and β-galactosidase as a reporter gene. P4-R5 cells are HeLa cells that contain a stably integrated β-galactosidase reporter gene under the control of the HIV1 LTR.

15

Materials:

- 1.) Cells: P4-R5 cells
2.) DNA: 0.78 µg/µL pcDNA3.1 zeo (+)APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695)
20 3.) Transfection reagents: FUGENE®
4.) Media: OPTIMEM®
cDMEM (-)phenol red /10% FBS
5.) Compounds: Compound X (β-secretase inhibitor) 10 mM
L-875,532 (γ-secretase inhibitor) 10 mM

25

Protocol:

Day 1

- 1.) Cell count on P4-R5 cells = 7.6×10^5 cells per ml in cDMEM (-)PR.
Seeded sterile white luminometer TC plates with the following cell numbers:
30 10 ml
 $5 \times 10^3/\text{well} = 0.75 \text{ ml}$ in 9.25 ml media
 $1.0 \times 10^4/\text{well} = 1.5 \text{ ml}$ in 8.5 ml media

Seeded 100 μ L cells per well.
Incubate overnight at 37°C, 5% CO₂.

Day 2

2.) Made up media with appropriate dilutions of inhibitors.

- 5 On no-inhibitor controls, added 100 μ L of cDMEM with 1% DMSO
 On wells with Compound X, added 10 μ M inhibitor in cDMEM
 On wells with L-875,532, added 10 μ M inhibitor in cDMEM
- 3.) Prior to transfection, pulled off media on P4-R5 cells and replaced with media -/+ inhibitor.
- 10 FUGENE® transfection:
 For FUGENE® transfection:
 4.) Added 600 μ L of OPTIMEM® to sterile EPPENDORF® tube and carefully added 30 μ L FUGENE® to media, without touching walls of tube. Incubated at room temperature for 5 minutes.
- 15 In separate EPPENDORF® tubes, added each DNA.
 Added FUGENE®/OPTIMEM® dropwise to DNA; incubated at room temperature for 15 minutes.
 Added 15 μ L /well of DNA/FUGENE®/OPTIMEM® dropwise to media in appropriate wells on P4-R5 cells, swirling to mix.
- 20

TABLE 4

Transfection number	Conc of DNA	Vol of DNA	Vol of FUGENE®	Vol of sterile OPTIMEM®
APP-ML-Tat-APPct	0.78 μ g/ μ L	5 μ g = 6.5 μ L	30 μ L of FUGENE®	600 μ L of OPTIMEM®
pUCd5TAT	1.24 μ g/ μ L	5 μ g = 4.0 μ L	30 μ L of FUGENE®	600 μ L of OPTIMEM®
p243-4	0.56 μ g/ μ L	5 μ g = 9 μ L	30 μ L of FUGENE®	600 μ L of OPTIMEM®

In Table 4, "APP-ML-Tat-APPct" refers to pcDNA3.1 zeo (+)APP(1-651)SW,
 25 K612V-TATexonI(M1L) APP(664-695). "pUCd5TAT" is an expression vector that

serves as a positive control for TAT expression, since it is a construct in which TAT is under the control of a strong, constitutive promoter (see Figure 24). "p243-4" is a control expression vector that directs the expression of APP.

- 5 5.) Plates were transferred to an incubator and incubated for 48 hours to allow expression and processing of the proteins.

Day 4

- 6.) The protocol below was followed for lysis of the cells and measurement of β -galactosidase in the cell lysates.

10

Measurement of β -galactosidase in lysates of transfected cells.

1. Removed TROPIX® chemiluminescence kit(s) from cold room, allowed to come to room temperature in a 37°C water bath.

2. β -galactosidase standards were prepared:

- 15 Made 1:5000 dilution of β -galactosidase stock (1 mg/ml) in lysis buffer. Did 2 fold dilutions.

3. Diluted TROPIX® substrate 1:25 into buffer. (Made enough for 100 μ L /well).

4. Added to reservoir and added 100 μ L/well.

- 20 5. Added 10 μ L of β -galactosidase standards to column 12 on plate and incubated in dark for 1 hour.

6. Read immediately in luminometer using standard file. Filled in required fields, read plate.

- 25 The results are shown in Figure 33. Figure 33 demonstrates that the present invention was able to identify both the β -secretase inhibitor (Compound X) and the γ -secretase inhibitor (L-875,532). In Figure 33, "APP-tat-ct32" refers to pcDNA3.1 zeo (+)APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695). Although not indicated in Figure 33, the results for the controls were as expected: a large transactivation of

- 30 the LTR by pUCd5STAT was observed which was not affected by either inhibitor. No transactivation was seen with p243-4.

EXAMPLE 9

Comparison of the use of APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695) and APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695) with a β -galactosidase reporter gene

5 The following shows a side-by-side comparison of the practice of the present invention with the APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695) fusion protein (SEQ ID NO:2) and the APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695) fusion protein (SEQ ID NO:4). P4-R5 cells are HeLa cells that contain a stably integrated β -galactosidase reporter gene under the control of the HIV1 LTR.

10

Materials:

1.) Cells: P4-R5 cells

2.) DNA: 0.78 μ g/ μ L pcDNA3.1 neo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695)

15 0.812 μ g/ μ L pcDNA3.1 neo (+) APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695)

 1.24 μ g/ μ L pUCd5TAT

 0.56 μ g/ μ L p243-4

3.) Transfection reagents: FUGENE®

20 4.) Media: OPTIMEM®

cDMEM (-)phenol red /10% FBS

5.) Compounds: Compound X (β -secretase inhibitor) 10 mM
 L-875,532 (γ -secretase inhibitor) 10 mM

Day 1

25 1.) Cell count on P4-R5 cells = 5×10^5 cells per ml in cDMEM (-)PR.

Seeded sterile white luminometer TC plates with the following cell numbers:

10 ml

5×10^3 /well = 4.0 ml in 36.0 ml media

Diluted 1:1 into media and seeded one plate at 2.5×10^3 /well.

30 Seeded 100 μ L cells per well.

Incubated overnight at 37°C, 5% CO₂.

Day 2

- 2.) Made up media with appropriate dilutions of inhibitors.
 On no-inhibitor controls, added 100 µL of cDMEM with 1% DMSO
 On wells with Compound X, added titration curve from 100 µM
 5 inhibitor in cDMEM (-)PR.
 On wells with L-875,532, added titration curve from 100 µM inhibitor
 in cDMEM (-)PR.
- 10 3.) Prior to transfection, pulled off media on P4-R5 cells and replaced with media -/+
 inhibitor.

FUGENE® transfection:

- 15 4.) Added volume of OPTIMEM® to sterile EPPENDORF® tube and carefully
 added correct volume of FUGENE® to media, without touching walls of tube.
 Incubated at room temperature for 5 minutes.
 In separate EPPENDORF® tubes, added each DNA, as outlined below.
 Added FUGENE®/OPTIMEM® dropwise to DNA; incubated at room
 temperature for 15 minutes.
- 20 Added 15 µL/well of DNA/FUGENE®/OPTIMEM® dropwise to media in
 appropriate wells on P4-R5 cells, swirling to mix.

TABLE 5

Transfection number	Conc of DNA	Vol of DNA	Vol of FUGENE®	Vol of sterile OPTIMEM®
1.) APP-ML-Tat-APPct (Sw)	0.78 µg/µL	10 µg = 13 µL	60 µL of FUGENE®	1200 µL of OPTIMEM®
2.) APP-ML-Tat-APPct (WT)	0.812 µg/µL	10 µg = 12.2 µL	60 µL of FUGENE®	1200 µL of OPTIMEM®
3.) pUCd5TAT	1.24 µg/µL	5 µg = 4.0 µL	30 µL of FUGENE®	600 µL of OPTIMEM®
4.) p243-4	0.56 µg/µL	5 µg = 9 µL	30 µL of FUGENE®	600 µL of OPTIMEM®

In Table 5, "APP-ML-Tat-APPct (Sw)" refers to pcDNA3.1 neo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695). "APP-ML-Tat-APPct (WT)" refers to pcDNA3.1 neo (+) APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695).

- 5 "pUCd5STAT" is an expression vector that serves as a positive control for TAT expression, since it is a construct in which TAT is under the control of a strong, constitutive promoter (see Figure 24). "p243-4" is a control expression vector that directs the expression of APP.
- 10 5.) Plates were transferred to an incubator and incubated for 36 hours to allow expression and processing of proteins.

Day 4

- 15 6.) The protocol below was followed for lysis of the cells and measurement of β -galactosidase in the cell lysates.

Measurement of β -galactosidase in lysates of transfected cells.

- 20 1. Removed TROPIX® chemiluminescence kit(s) from cold room, allowed to come to room temperature in a 37°C water bath.
2. β -galactosidase standards were prepared:
Made 1:5000 dilution of β -galactosidase stock (1 mg/ml) in lysis buffer. Did 2 fold dilutions.
- 25 3. Diluted TROPIX® substrate 1:25 into buffer. (Made enough for 120 μ L /well).
4. Added to reservoir and added 120 μ L/well.
5. Added 10 μ L of β -galactosidase standards to column 12 on plate and incubated in dark for 1 hour.
- 30 6. Read immediately in luminometer using standard file. Filled in required fields, read plate.

The results are shown in Figure 34. In Figure 34, "APP(NFEV)HAMycFLAG" refers to a protein that is a variant of APP in which NFEV is present at the β -secretase cleavage site and there are epitope tags in the amino terminal portion of the protein

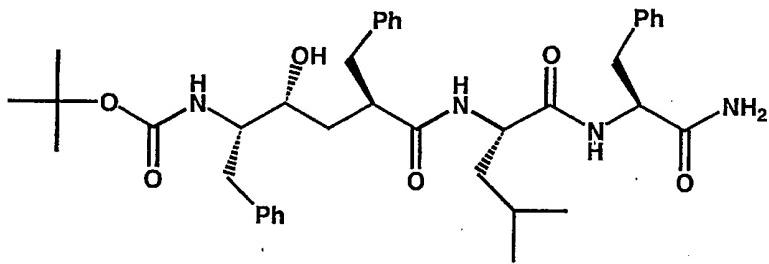
but there is no transcription factor fused to APP. "APP(Sw)tat-ct32" refers to pcDNA3.1 neo (+) APP(1-651)SW, K612V-TATexonI(M1L) APP(664-695). "APP(WT)tat-ct32" refers to pcDNA3.1 neo (+) APP(1-651)wt, K612V-TATexonI(M1L) APP(664-695). Figure 34 shows that the Swedish version and the wild-type version of APP appear to work about equally well in the assay.

EXAMPLE 10

L-685,458

10

L-685,458 is a γ -secretase inhibitor having the following structure:



L-685,458 contains an hydroxyethylene dipeptide isostere and is thought to function as a transition state analog mimic of aspartyl proteases (Shearman et al., 2000, Biochemistry 39:8698-8704). L-685,458 was prepared as follows:

- 15 {1S-Benzyl-4R-[1-(1S-carbamoyl-2-phenylethylcarbamoyl)-1S-3-methylbutylcarbamoyl]-2R-hydroxy-5-phenylpentyl}carbamic acid tert-butyl ester (L-685,458) was prepared by the coupling of 2R-benzyl-5S-tert-butoxycarbonylamino-4R-(tert-butyldimethylsilanyloxy)-6-phenylhexanoic acid (Evans et al., 1985, J. Org. Chem. 50:4615-4625) with Leu-Phe-NH₂ followed by
20 deprotection with tetrabutylammonium fluoride. The synthesis of {1S-benzyl-4R-[1-(1S-carbamoyl-2-phenylethylcarbamoyl)-1S-3-methylbutylcarbamoyl]-2S-hydroxy-5-phenylpentyl}carbamic acid tert-butyl ester (L-682,679) has been described previously (De Solms et al., 1991, J. Med. Chem. 34:2852-2857). {1S-Benzyl-4R-[1-(1S-carbamoyl-2-phenylethylcarbamoyl)-1S-3-methylbutylcarbamoyl]-2-oxo-5-phenylpentyl}carbamic acid tert-butyl ester (L-684,414) was prepared by pyridinium dichromate-mediated oxidation of L-682,679.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from 5 the foregoing description. Such modifications are intended to fall within the scope of the appended claims.

Various publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

WHAT IS CLAIMED IS:

1. A DNA molecule comprising a nucleotide sequence encoding a fusion protein comprising amino acids 589-651 selected from the group consisting of
5 wild type APP695, the Swedish version of APP695 and the NFEV (SEQ ID NO:40) version of APP695 and a transcription factor where the transcription factor is fused in frame to the carboxyl terminus of amino acids 589-651.
2. The DNA molecule of claim 1 where amino acids 589-651
10 contain a K612V mutation.
3. The DNA molecule of claim 1 where the nucleotide sequence further encodes amino acids 664-695 of APP695 wherein amino acids 664-695 are fused in frame to the carboxyl terminus of the transcription factor.
15
4. The DNA molecule of claim 1 where the transcription factor is selected from the group consisting of: HIV-1 TAT, Gal4-VP16, the entire Gal4 protein, LexA-VP16, E12, E47, Twist, Papillomavirus E2, EBV Zta, BIV TAT, HIV-2 TAT, or SIV TAT.
20
5. The DNA molecule of claim 3 where the fusion protein is selected from the group consisting of: APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:4); APP(1-651)wt, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:8); APP(1-651)wt, TATexonI(M1L) APP (664-695) (SEQ ID NO:12);
25 and APP(1-651)wt, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:16).
6. The DNA molecule of claim 3 where the fusion protein is selected from the group consisting of: APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:2); APP(1-651)SW, K612V, Gal4-VP16(M1L) APP (664-
30 695) (SEQ ID NO:6); APP(1-651)SW, TATexonI(M1L) APP (664-695) (SEQ ID NO:10); and APP(1-651)SW, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:14).
7. The DNA molecule of claim 3 where the fusion protein is selected from the group consisting of: APP(1-651)NFEV, K612V-TATexonI(M1L)

APP (664-695) (SEQ ID NO:23) and APP(1-651)NFEV, K612V, GAL4-VP16(M1L)
APP (664-695) (SEQ ID NO:25).

8. An expression vector comprising the DNA molecule of claim 1.
- 5
9. A eukaryotic cell comprising the DNA molecule of claim 1.
10. The cell of claim 9 further comprising a reporter gene where the reporter gene is under the control of a regulatory DNA sequence that is capable of being activated by the transcription factor.
11. A method of identifying a substance that inhibits APP processing comprising:
 - (a) providing a recombinant eukaryotic cell which:
 - 15 (i) expresses a fusion protein comprising amino acids 589-651 selected from the group consisting of wild type APP695, the Swedish version of APP695 and the NFEV (SEQ ID NO:40) version of APP695 and a transcription factor where the transcription factor is fused in frame to the carboxyl terminus of amino acids 589-651; and
 - 20 (ii) comprises a reporter gene operably linked to a regulatory DNA sequence which capable of being activated by the transcription factor;
 - (b) measuring the level of reporter gene product in the cell in the absence of the substance;
 - 25 (c) adding the compound to the cell and measuring the level of reporter gene product in the cell in the presence of the substance; where a decrease in the level of reporter gene product in the presence as compared to the absence of the substance indicates that the substance inhibits APP processing.
 - 30 12. The method of claim 11 where amino acids 589-651 contain a K612V mutation.
 13. The method of claim 11 where the transcription factor is selected from the group consisting of: HIV-1 TAT, Gal4-VP16, the entire Gal4

protein, LexA-VP16, E12, E47, Twist, Papillomavirus E2, EBV Zta, BIV TAT, HIV-2 TAT, or SIV TAT.

14. The method of claim 11 where the fusion protein is selected
5 from the group consisting of: APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:4); APP(1-651)wt, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:8); APP(1-651)wt, TATexonI(M1L) APP (664-695) (SEQ ID NO:12); and APP(1-651)wt, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:16).

10 15. The method of claim 11 where the fusion protein is selected from the group consisting of: APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:2); APP(1-651)SW, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:6); APP(1-651)SW, TATexonI(M1L) APP (664-695) (SEQ ID NO:10); and APP(1-651)SW, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:14).

15 16. The method of claim 11 where the fusion protein is selected from the group consisting of: APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695) (SEQ ID NO:23) and APP(1-651)NFEV, K612V, Gal4-VP16(M1L) APP (664-695) (SEQ ID NO:25).

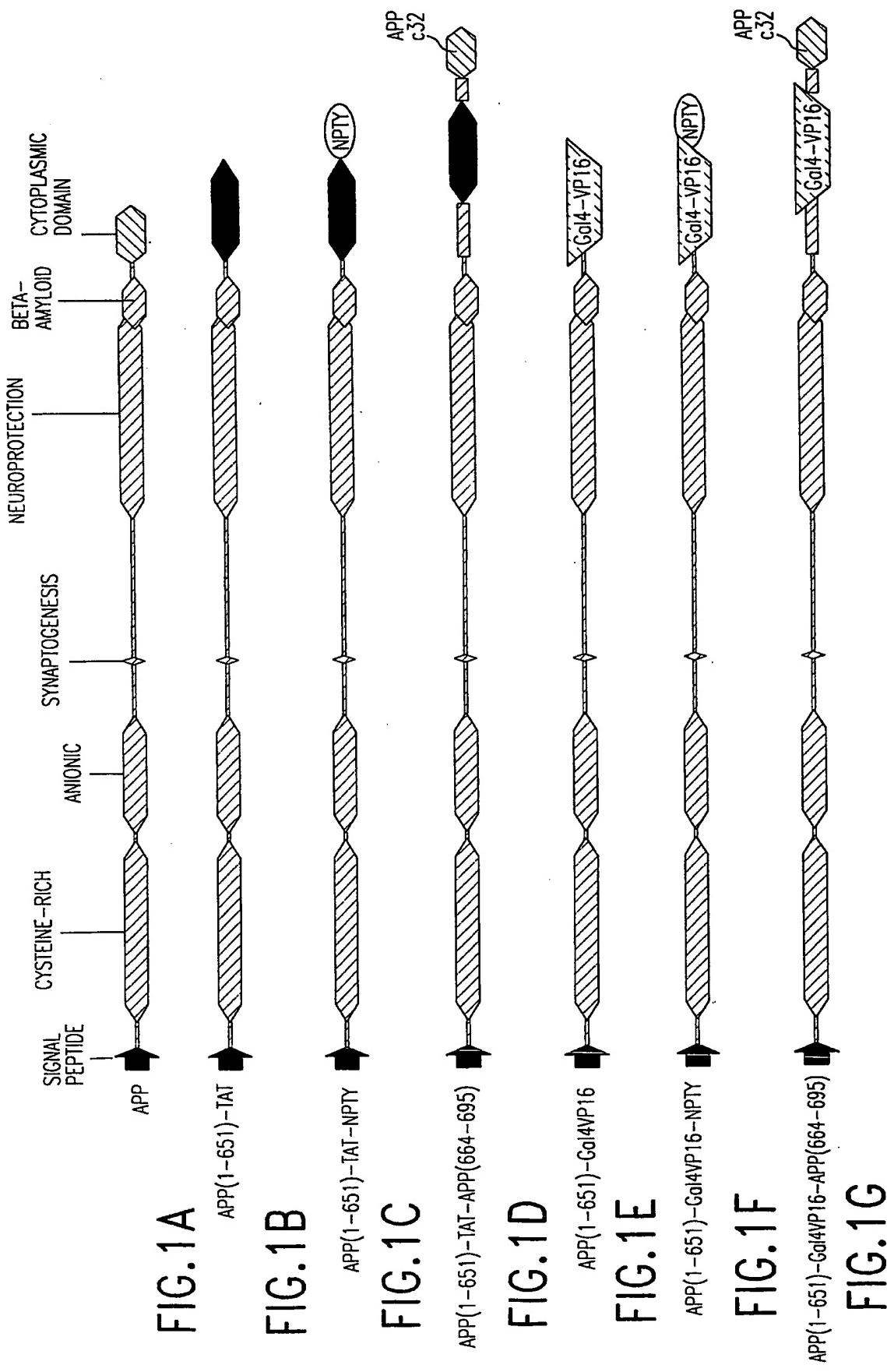
20 17. A method of identifying a substance that inhibits APP processing comprising:
25 (a) providing a recombinant eukaryotic cell which:
 (i) expresses a fusion protein comprising an amino acid sequence from APP that is capable of being cleaved by both β -secretase and γ -secretase and a transcription factor where the transcription factor is fused in frame to the amino acid sequence from APP; and
 (ii) comprises a reporter gene operably linked to a regulatory DNA sequence which capable of being activated by the transcription factor;
30 (b) measuring the level of reporter gene product in the cell in the absence of the substance;
 (c) adding the compound to the cell and measuring the level of reporter gene product in the cell in the presence of the substance;

where a decrease in the level of reporter gene product in the presence as compared to the absence of the substance indicates that the substance inhibits APP processing.

5 18. The method of claim 17 where the amino acid sequence from APP comprises an amino acid sequence selected from the group consisting of 589-651 of APP695, 589-651 of the Swedish version of APP695, and 589-651 of the NFEV version of APP695.

10 19. The method of claim 17 where the amino acid sequence from APP contains the amino acid sequence NLDA (SEQ ID NO:38) at the β -secretase cleavage site instead of the wild-type sequence KMDA (SEQ ID NO:34).

15 20. The method of claim 17 where the amino acid sequence from APP contains the amino acid sequence NFEV (SEQ ID NO:40) at the β -secretase cleavage site instead of the wild-type sequence KMDA (SEQ ID NO:34).



2/76

DNA sequence of APP(1-651)SW, K612V-TATexon1(M1 L) APP (664-695)
(SEQ ID NO: 1)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACCTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTTC CCTACCGCTG
351 CTTAGTTGGT GAGTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGCTGGTGG GGCGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACACC ACCATTGCCA CCACCACAC CACCACACA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCTCCT CGGCCTCGTC ACGTGTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGACAGA CAGCACACCC
1301 TAAAGCATTG CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

FIG.2A

3/76

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCA GTGGCC GAGGAGATT
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAATCTAGA TGCAGAATT
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGCG
1901 GTGTTGTAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATT
2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC
2051 CAGGAAGTCA GCCTAAAAC GCTTGTACCA ATTGCTATTG TAAAAAGTGT
2101 TGCTTCATT GCCAAGTTG TTTCATGACA AAAGCCTTAG GCATCTCTA
2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA
2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC
2251 GCAGACGCCG CTGTCACCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA
2301 GAACGGCTAC GAAAATCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT
2351 AG

FIG.2B

4/76

(SEQ ID NO: 2)

Amino acid sequence of APP(1-651)SW, K612V-TATexonI(M1L) APP (664-695)

m1pglal111aawtaralevptdgnag11aepqiamfcgr1nmhmnnvqngkwdsdpsgtktcidtkegilqycq
evyapelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckflhqermdvcethlh
whtvaketcsekstn1hdygml1pcgidkfrgvefvccplaeesdnvdssadaeeddsdvwggadtdyadgs
1 edkvvevaeeeevaeveeeeadddeddedgeveeeeapeyeaterttsiatttttesveevrvpttaastpd
avdkyletpgdenehahfqkakerleakhrermssqvmoreeweaerqaknlpkadkkaviqhfkqekvesleqe
aanerqq1vethmarveamndrrrlaleniyitalqavpprprhvfnmlkkyvraeqkdrqht1khfehvrvmvd
pkkaaqirsqvmtth1rviyermnqs1s1lynvpavaeeiqdevde11qkeqnysddvlammisepri sygnda1
mps1tetkttvel1pvngefs1dd1qpwhsfadsvpantenevevpdarpaadrg1ttrpgsg1tnikteeisev
2 3 4 5
n1daefrhdsgyevhhqy1vffaedvgsnkaiiglmvggyviatyivit1vm1kkk1gtelgstspwms
6
ad1lepvdpr1lepwkhpqsgpktac1cykkccfhcqvcfm1kalgiisygrkkrrrrahqnsqthgas1skg
7 8
r1s1stvaaadaavtpeerh1skmqqngnyenptykffeqmgn

FIG.3

5/76

DNA sequence of APP(1-651)wt, K612V-TATexon1(M1L) APP (664-695)
(SEQ ID NO: 3)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACCTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTG CCTACCGCTG
351 CTTAGTTGGT GAGTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGCTGGTGG GGCAGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTACA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACACC ACCATTGCCA CCACCAACAC CACCACACCA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTT CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCTCCT CGGCCTCGTC ACGTGTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGACAGA CAGCACACCC
1301 TAAAGCATTG CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTATG AGCGCATGAA

FIG.4A

6/76

1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAGTGGCC GAGGAGATT
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGG AGTCTCTGAAG TGAAGATGGA TGCAGAATT
1801 CGACATGACT CAGGATATGA AGTCATCAT CAAGTATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATT
2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAAGCATC
2051 CAGGAAGTCA GCCTAAAAC GCTTGTACCA ATTGCTATTG TAAAAAGTGT
2101 TGCTTTCATT GCCAAGTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA
2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA
2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC
2251 GCAGACGCCG CTGTCACCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA
2301 GAACGGCTAC GAAAATCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT
2351 AG

FIG.4B

7/76

(SEQ ID NO: 4)

Amino acid sequence of APP(1-651)wt, K612V-TATexonI(M1L) APP (664-695)

m₁pglal₁₁₁laawtaralevptdgnag₁₁aepqiamfcgrl_{nmhm}nvqngkwdsapsgtktcidtk_{egil}qycq
evypelqitn_vveanqpvtiqnwckrgrkqckthphfvipyrc₁vgefis_dallvpdkckf₁hqermdvceth₁h
whtvaketcsekstn₁hdyg_{ml}pcgidkfrgvefvccp₁aeesdnvd_{sadae}eddsdvwwggadtdyadgs
edkvvevaeee_evaeveeeeadddedgedveeeeaeypyeeaterttsiattttttesveevrvptaastpd
avdkyletpgdenehahfqkakerleakhrermsqv_mrew_{ea}erqakn_{lp}kadkkaviqh_fqekvesleqe
aanerqq₁vethmarveam_{ndrrr}1alenytalqavpprprhvfnm_{lkky}vraeqkdrqht₁khfehvrmvd
pkkaaqirsqvmth₁rviyermnqs₁lynpavaeei_qdevde_{ll}qkeqnysddv_{lan}misepri_{sy}gndal
mps₁tetkttv_{ell}pvngefs₁dd₁qpwhs_{fgad}svpantenevepvdarpaadrg₁ttrpgs_{gl}tniktee_{isev}
2₃₄₅
kmdaefrhdsgyevhhq_{vl}vfaedvgsnk_{gai}ig_{lmvggyvi}atv_{ivitlv}m₁kkkk_{lgte}gstsp_{wwns}
6
adilepvdp₁lepwk_{hpq}sqpkta_ctncy_{ckcc}fhc_{qcvcf}mtkalq_{is}ygrkrrqrrrahqnsqthqaslska
7₈
risstvaaadaavtpeerh₁skmqngyenptykffeqmgn

FIG.5

8/76

DNA sequence of APP(1-651)SW, K612V-GAL4VP16(M1 L) APP (664-695)
(SEQ ID NO: 5)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAAGTCTA CCCTGAACTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTTC CCTACCGCTG
351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGATGAA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCAGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTACA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACCAACC AGCATTGCCA CCACCAACAC CACCAACACA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCCT CGGCCTCGTC ACGTGTTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGACAGA CAGCACACCC

FIG.6A

9/76

1301 TAAAGCATT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCA GTGGCC GAGGAGATT
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGGAG ATCTCTGAAG TGAATCTAGA TGCAGAATT
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATT
2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTGCC
2051 GACTAAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT
2101 CTGAAGAACAA ACTGGGAGTG TCGCTACTCT CCCAAACCCA AAAGGTCTCC
2151 GCTGACTAGG GCACATCTGA CAGAAGTGGAA ATCAAGGCTA GAAAGACTGG
2201 AACAGCTATT TCTACTGATT TTTCTCGAG AAGACCTTGA CATGATTTG
2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTGT
2301 ACAAGATAAT GTGAATAAAG ATGCCGTAC AGATAGATTG GCTTCAGTGG
2351 AGACTGATAT GCCTCTAACAA TTGAGACAGC ATAGAATAAG TGCGACATCA
2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG
2451 AATTCCCGGG GATCTGGCCC CCCCCACCGA TGTCAGCCTG GGGGACGAGC
2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC
2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG. 6B

10/76

2601 TACCCCCCAC GACTCCGCC CCTACGGCGC TCTGGATATG GCCGACTTCG
2651 AGTTTGAGCA GATGTTTACC GATGCCTTG GAATTGACGA GTACGGTGGG
2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG
2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT
2801 TCTTTGAGCA GATGCAGAAC TAG

FIG.6C

11/76

(SEQ ID NO: 6)

Amino acid sequence of APP(1-651)SW, K612V, GAL4-VP16(deIM1) APP (664-695)

m1pg1all11laawtaralevptdgnag11aepqiamfcgr1nmhmnvqngkwdspsgtktcidtkegilqycq
evyapelqitnvveanqvptiqnwckrgrkqckthphfvipyrc1vgefis الد lvpdkckf1hqeरmdvceth1h
whtvaketcsekstn1hdygm1lpcgidkfrgvefvccplaeesdnvsadaeeddsdvwggadtdyadgs
1
edkvvevaeeeevaeveeeeadddeddedgeveeeaeepyeeaterttsiattttttesveevrvpttaastpd
avdkyletpgdenehahfqkakerleakhrermisqvmreweeaerqaknlpkadkkaviqhfkqekvesleqe
aanerqq1vethmarveam1ndrrrlaleniyitalqavpprprhvfnn1kkyvraeqkdrqht1khfehvrnvd
pkkaaqirsqvmth1rviyermnqs1s1lynvpavaeeiqdevde11qkeqnysddvlamiseprisygnda1
mps1tetkttvellpvngefs1dd1qpwhsfadsvpantenevevpdarpaadrg1ttrpgsg1tnikteeisev
2 3 4 5
nldaefrhdsgyevhhqylvffaedvgsnkgaiglmvggyviatvivit1vmlkkkk1gtelgstspwwns
6
adi1k1ssieqacd1c1kk1kcskekpkcakc1knncryspktkrsp1trah1tevesrlrerleg1f11ifpred1d
milkmndlqdikalltg1fvqdvnkdavtdrlasvetdmp1t1rqhrisatssseesssnkqqrqltvsgipqdlapp
7 8
tdvslqde1h1dqgedvamahadalddfd1dm1qdqdspqpgftphdsapyqaldmadfefeqmftdalqidey
ggdiqhsaaaadaavtpeerh1skmqngnyenptykffeqmgn

FIG. 7

12/76

DNA sequence of APP(1-651)wt, K612V, GAL4-VP16(deIM1) APP (664-695)
(SEQ ID NO: 7)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAAC TG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATT CCTACCGCTG
351 CTTAGTTGGT GAGTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTACA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACCAACC AGCATTGCCA CCACCAACAC CACCAACACA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCCTCGTC CGGCCTCGTC ACGTGTTCAA

FIG.8A

13/76

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGGACAGA CAGCACACCC
1301 TAAAGCATTG CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCA GTGGGCC GAGGAGATTG
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAAAGT TTCAGATGAC
1501 GTCTTGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAAGATGGA TGCAGAATTG
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGCG
1901 GTGTTGTAT AGCGACAGTG ATCGTCATCA CCTTGGTGT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTG
2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTGCG
2051 GACTAAAAAA GCTCAAGTGC TCCAAAGAAA ACCGAAGTG CGCCAAGTGT
2101 CTGAAGAACAA ACTGGGAGTG TCGCTACTCT CCCAAACCA AAAGGTCTCC
2151 GCTGACTAGG GCACATCTGA CAGAAGTGGATCAAGGCTA GAAAGACTGG
2201 AACAGCTATT TCTACTGATT TTTCTCGAG AAGACCTTGA CATGATTTG
2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT
2301 ACAAGATAAT GTGAATAAG ATGCCGTAC AGATAGATTG GCTTCAGTGG
2351 AGACTGATAT GCCTCTAACAA TTGAGACAGC ATAGAATAAG TGCGACATCA
2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG
2451 AATTCCCGGG GATCTGGCCC CCCCCGACCGA TGTCAGCCTG GGGGACGAGC
2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC
2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.8B

14/76

2601 TACCCCCCAC GACTCCGCC CCTACGGCGC TCTGGATATG GCCGACTTCG
2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG
2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG
2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT
2801 TCTTGAGCA GATGCAGAAC TAG

FIG.8C

15/76

(SEQ ID NO: 8)

Amino acid sequence of APP(1-651)wt, K612V, GAL4-VP16(de1M1) APP (664-695)

m1pg1a111aawtaralevptdgnag1laepqiamfcgrlnmhmnvqngkwdspsgtktcidtkegilqycq
evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisda11vpdkckf1hqeरmdvceth1h
whtvaketcsekstn1hdym11pcgidkfrgvefvccplaeesdnvdsadaeeddsdvwggadtdyadgs
1
edkvvevaeeeavaeveeeeadddedgedeveeeaeeyeeaterttsiatttttesveevrvpttaastpd
avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqakn1pkadkkaviqhfkqekvesleqe
aanerqq1vethmarveam1ndrrrlaleniyitalqavpprprhvfnmilkkyvraeqkdrqht1khfehvrmvd
pkkaaqirsqvmtlrviyermnqs1s1lynvpavaeeiqdevde11qkeqnysddvlanmisseprisyndal
mps1tetkttvel1pvngefs1dd1qpwhsfadsvpanterevepvdarpaadrg1ttrpgsg1tnikteeisev
2 3 4 5
kmdaefrhdsgyevhhqk1vfaedvgsnkgaig1mvggviatvivit1vm1kkkk1gtelgstspwns
adik11ssiegacd1crlkk1kskekpkcakc1knncryspktkrsp1trahltevesrler1eq1f11ifpred1d
6
milkm1s1qdik11tq1fvqdvnkdavtdrlasvetdmpl1rqhrisatssseessnkqqr1tvsgipgd1app
tdvs1gde1h1dgedvamahad1ddfd1dm1gdgdspgpgftphdsapygaldmadfefeqmf1dalqidey
7 8
ggdiqhsaaaadaavtpeerh1skmqnqyenptykffeqmgn

FIG.9

16/76

DNA sequence of APP(1-651)SW, TATexon1(M1 L) APP (664-695)

(SEQ ID NO: 9)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAAC CCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAAC TG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATT CCTACCGCTG
351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGGAGG ATGACTCGGA TGTCTGGTGG GGCAGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTACA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACCAACC AGCATTGCCA CCACCAACAC CACCAACACA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCCTCGT CGGCCTCGTC ACGTGTTCAA

FIG. 10A

17/76

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGCACAGA CAGCACACCC
1301 TAAAGCATT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCACTGGCC GAGGAGATTG
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAATCTAGA TGCAGAATTG
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG
1901 GTGTTGTCA AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTG
2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC
2051 CAGGAAGTCA GCCTAAAAC GCTTGTACCA ATTGCTATTG TAAAAGTGT
2101 TGCTTCATT GCCAAGTTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA
2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA
2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC
2251 GCAGACGCCG CTGTCACCCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA
2301 GAACGGCTAC GAAAATCCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT
2351 AG

FIG.10B

18/76

(SEQ ID NO: 10)

Amino acid sequence of APP(1-651)SW, TATexonI(M1L) APP (664-695)

m1pglal111aawtaralevptdgnag11aepqiamfcgrlnmhmnvqngkwdsdpsgtktcidtkegilqycq
evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrc1vgefisdal1vpdkckf1hqermvdvceth1h
whtvaketcsekstn1hdyg11pcgidkfrgvefvccplaeesdnvsadaeeddsdvwggadtdyadgs
1
edkvvevaeeeevaeveeeeadddeddedgdeveeeaeypyeeaterttsiattttttesveevrvptaastpd
avdkyletpgdenehahfqakerleakhrermsqvwmreweeaerqakn1pkadkkaviqhfkvesleqe
aanerqq1vethmarveamndrrrlalenytalqavpprprhvfnm1kkyvraeqkdrqht1khfehvrmd
pkkaaqirsqvmth1rviyermnqs1s1lynvpavaeeiqdevde11qkeqnyssddvlammiseprisyndal
mps1tetkttvellpvngefs1dd1qpwhsfadsvpantenevepvdarpaadrglttrpgsg1tnikteeisev
2 3 4 5
nldaefrhdsgyevhhqk1vffaedvgsnkgaig1mvggvviatyivit1vmlkkkk1gtelgstspwvns
6
ad1lepvdp1lepwkhpqsqpktaactncycckccfhcqvcf1kalqisygrkkrrrrahqnsqthgas1skq
7 8
risstvaaadaavtpeerh1skmqngyenptykffeqmn

FIG.11

19/76

DNA sequence of APP(1-651)wt, TATexon1(M1L)-APP (664-695)

(SEQ ID NO: 11)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCAGAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTTC CCTACCGCTG
351 CTTAGTTGGT GAGTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCAGTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCAGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTACA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACACC ACCATTGCCA CCACCAAC CACCACACA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAAATCTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCCT CGGCCTCGTC ACGTGTCAA

FIG.12A

20/76

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAGA GAAGGGACAGA CAGCACACCC
1301 TAAAGCATTT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCA GTGGGCC GAGGAGATTG
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAAAGT TTCAGATGAC
1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAAGATGGA TGCAGAATTG
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG
1901 GTGTTGTATC AGCGACAGTG ATCGTCATCA CCTTGGGTGAT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATTG
2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATT
2051 CAGGAAGTCA GCCTAAAATC GCTTGTACCA ATTGCTATTG TAAAAAGTGT
2101 TGCTTTCATT GCCAAGTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA
2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA
2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC
2251 GCAGACGCCG CTGTCACCCCC AGAGGGAGCGC CACCTGTCCA AGATGCAGCA
2301 GAACGGCTAC GAAAATCCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT
2351 AG

FIG. 12B

21/76

(SEQ ID NO: 12)

Amino acid sequence of APP(1-651)wt, TATexonI(M1L) APP (664-695)

m₁pglal₁₁₁laawtaralevptdgnag₁₁aepqiamfcgr₁n_mhmnvqngkwdsdpsgtktcidtkegilqycq
evypelqitnveanqpvtiqnwckrgrkqckthphfvipyrc₁vgefis_dallvpdkckf₁hqermdvceth₁h
whtvaketcsekstn₁hdyg₁₁pcgidkfrgvefvccplaeesdnvd_sadaeeddsdvwggadtdyadgs
edkvvevaeeeavaeveeeeadddeddedgeveeeaeypyeeaterttsiattttttesveevrvpttaastpd
avdkyletpgdenehahfqkakerleakhrermsqv_mrew_eeaerqakn₁pkadkkaviqh_fqekvesleqe
aanerqqlvethmarveam₁ndrrrlalenytitalqavpprprhvfnm₁kkyvraeqkdrqht₁khfehvrmvd
pkkaaqirsqv_mth₁rviyer_mnqs₁s₁₁ynpavaeeiqdevde₁₁qkeqnysdd₁anm_iseprisygnd₁
mpsltetkttve₁₁pvngefs₁dd₁qpwhsf_gadsvpantenevepvdarpaadrg₁ttrpgsg₁tnikteeisev
2₂3₃4₄5₅
kmdaefrhsgyevhhqk₁vffaedvgsnk_gaiig₁mvggvviatvivit₁vm₁kkk₁getelg₁stsp_wwns
6₆
adilepvdpriepwkhpqspktactncyc₁kkccfhcqvcfm_tkalqisyqrkrrrrahqnsqthqas₁skq
7₇8₈
rissstvaaadaavtpeerhlskmqqngyenptykffeqmgn

FIG.13

22/76

DNA sequence of APP(1-651)SW, GAL4VP16(de1Met) APP (664-695)

(SEQ ID NO: 13)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTTC CCTACCGCTG
351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGATGCA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCGGAGCAG ACACAGACTA
651 TGCAAGATGGG AGTGAAGACA AAGTAGTACA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACCAACC AGCATTGCCA CCACCAACAC CACCAACACA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AACCAAAGAA CTTGCCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCCT CGGCCTCGTC ACGTGTTCAA

FIG.14A

23/76

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGGACAGA CAGCACACCC
1301 TAAAGCATT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCA GTGGCC GAGGAGATT
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGGCCA ACATGAATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAATCTAGA TGCAGAATT
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATT
2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTTGCC
2051 GACTAAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT
2101 CTGAAGAACAA ACTGGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC
2151 GCTGACTAGG GCACATCTGA CAGAAGTGGAA ATCAAGGCTA GAAAGACTGG
2201 AACAGCTATT TCTACTGATT TTTCTCGAG AAGACCTTGA CATGATTTG
2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT
2301 ACAAGATAAT GTGAATAAG ATGCCGTAC AGATAGATTG GCTTCAGTGG
2351 AGACTGATAT GCCTCTAACAA TTGAGACAGC ATAGAATAAG TGCGACATCA
2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG
2451 AATTCCCGGG GATCTGGCCC CCCCCACCGA TGTCAGCCTG GGGGACGAGC
2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC
2551 GATTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCGGGATT

FIG.14B

24/76

2601 TACCCCCCAC GACTCCGCC CCTACGGCGC TCTGGATATG GCCGACTTCG
2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG
2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG
2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT
2801 TCTTGAGCA GATGCAGAAC TAG

FIG.14C

25/76

(SEQ ID NO: 14)

Amino acid sequence of APP(1-651)SW, K612V, GAL4-VP16(de1M1) APP (664-695)

m₁pglal111aawtaralevptdgnagllae₂pqiamfcgrl₃nmmnvnqngkwdsdpsgtktcidtkegilqycq
evy₄pelqitnvveanqpvtiqnwckrgrkqckthphfvipyrc1vgefis₅dallvpdkckflhqermdvcethlh
whtvaketcsekstn₆hdym₇lpcgidkfrgvefvccplaeesdnvd₈sadaeeddsdvwwggadtdyadgs
edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiatttttesveevrvptaastpd
avdkyletpgdenehahf₉fqkakerleakhrermsqv₁₀rew₁₁eeaerqakn₁₂pkadkkaviqh₁₃f₁₄qekvesleqe
aanerqq₁₅lvet₁₆hmarveam₁₇ndrrrlalenitalqavpprprhvfnm₁₈lkkvraeqkdrqhtlk₁₉hfehvr₂₀mv₂₁d
pkkaaqirsqv₂₂mlrviyerm₂₃mnqs₂₄ls₂₅lynvpavaeeiqdevdellqkeqnysddv₂₆lanm₂₇iseprisygnda₂₈l
mps₂₉ltetkttve₃₀lpvngefs₃₁ddlqpwhsf₃₂gads₃₃vpantene₃₄evpvdarpaadrg₃₅l₃₆trpgsg₃₇l₃₈triktee₃₉isev
2 3 4 5
nldaefrhdsgyevhhq₄₀l₄₁vfaedvgsnk₄₂gaiigl₄₃mv₄₄gvviatv₄₅ivitl₄₆vm₄₇lkkk₄₈l₄₉telg₅₀stsp₅₁wns
ad₅₂ik₅₃lssieqacd₅₄dicrl₅₅kk₅₆l₅₇kcskekpkcak₅₈c₅₉kn₆₀we₆₁cr₆₂ysp₆₃ktkr₆₄sp₆₅trah₆₆ite₆₇ves₆₈r₆₉ler₇₀leg₇₁lf₇₂if₇₃pred₇₄ld
6
milkmdslqdikalltg1fvqdnvkdavtdrlasvetdmpl1rqhrisatssseesssnkgqr1tvsgipqdlapp
tdvs1gde1hldgedvamahada1ddf1dm1gdgdpgpqftphdsapygaldmadfefeqmftdalqidey
7 8
ggdiqhs₅₅gaaadaavtpeerh₅₆l₅₇sk₅₈mq₅₉ngyenptyk₆₀ffeq₆₁mqn

FIG. 15

26/76

DNA sequence of APP(1-651)wt, GAL4VP16(de1Met) APP (664-695)
(SEQ ID NO: 15)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACCTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTTC CCTACCGCTG
351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCAGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACACCAC AGCATTGCCA CCACCAACAC CACCAACACA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCCT CGGCCTCGTC ACGTGTTCAA

FIG.16A

27/76

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGACAGA CAGCACACCC
1301 TAAAGCATTG CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCA GTGGCC GAGGAGATT
1451 AGGATGAAGT TGATGAGCTG CTTAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCTTCCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAAGATGGA TGCAGAAATT
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATT
2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTGCC
2051 GACTTAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT
2101 CTGAAGAACAA ACTGGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC
2151 GCTGACTAGG GCACATCTGA CAGAAGTGGAA ATCAAGGCTA GAAAGACTGG
2201 AACAGCTATT TCTACTGATT TTTCCTCGAG AAGACCTTGA CATGATTTG
2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT
2301 ACAAGATAAT GTGAATAAAG ATGCCGTAC AGATAGATTG GCTTCAGTGG
2351 AGACTGATAT GCCTCTAACAA TTGAGACAGC ATAGAATAAG TGCGACATCA
2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG
2451 ATTCCCGGG GATCTGGCCC CCCCCGACCGA TGTCAGCCTG GGGGACGAGC
2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC
2551 GATTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.16B

28/76

2601 TACCCCCCAC GACTCCGCC CCTACGGCGC TCTGGATATG GCCGACTTCG
2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG
2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGGAGCG
2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT
2801 TCTTGAGCA GATGCAGAAC TAG

FIG.16C

29/76

(SEQ ID NO: 16)

Amino acid sequence of APP(1-651)wt, K612V, GAL4-VP16(deTM1) APP (664-695)

m1pg1all11aawtaralevptdgnag11aepqiamfcgrlnmhmnvqngkwdspsgtktcidtkegilqycq
evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisda11vpdkckf1hqermvdvceth1h
whtvaketcsekstn1hdym11pcgidkfrgvefvccplaeesdnvdsadaeeddsdvwggadtdyadgs
1
edkvvevaeeeavaeveeeeadddedgedgeveeeeaeypyeeaterttsiatttttesveevrvpttaastpd
avdkyletpgdenehahfqkakerleakhrermsqvmreweeaerqakn1pkadkkaviqhfkqekves1eqe
aanerqq1vethmarveam1ndrrrlaleniyitalqavpprprhvfnmilkkyvraeqkdrqht1khfehvrmvd
pkkaaqirsqvmtlrviyermnqs1s1lynvpavaeeiqdevde11qkeqnysddvlancmiseprisyndal
mps1tetkttve11pvngefs1dd1qpwhsfagsvpantenevepvdarpaadrg1ttrpgsg1tnikteeisev
2 3 4 5
kmdaefrhsgyevhhqk1vfaedvgsnkaiig1mvggviatvivit1vm1kkkk1gtelgstspwwwns
adik1ssiegacd1crlkk1kcskekpkcakc1knncryspktkrsp1trahltevesrler1eq1f11ifpred1d
6
milkmds1qdikalltg1fvqdnvnkdavtdrlasvetdmpltlrqhrisatssseessnkqqrqltvsqipqdlapp
tdvs1qde1h1dqedvamahad1ddfd1dm1qdqdspapgfphdsapygaldmadfefeqmftdalgidey
7 8
ggdiqhsgaaaadaavtpeerh1skmqqngyenptykffecmgn

FIG. 17

30/76

(SEQ ID NO: 17)

1 agtttcctcg gcagcggtag gcgagagcac gcggaggagc gtgcgcgggg gccccgggag
 61 acggcggcg tggcggcg ggcagagcaa ggacgcggcg gatccactc gcacagcagc
 121 gcactcggtg ccccgccgca ggtcgcgatg ctgcccggtt tggactgct cctgctggcc
 181 gcctggacgg ctcgggcgtt ggaggtaccc actgatggta atgctggcct gctggctgaa
 241 ccccagattt ccatgttctg tggcagactg aacatgcaca tgaatgtcca gaatgggaag
 301 tgggattcag atccatcagg gaccaaaacc tgcattgata ccaaggaagg catcctgcag
 361 tattgccaaag aagtctaccc tgaactgcag atcacaatg tggtagaagc caaccaacca
 421 gtgaccatcc agaactggtg caagcgggc cgcaagcagt gcaagaccca tccccacttt
 481 gtgattccct accgctgctt agttggtagg tttgttaatg atgcccctt cgttcctgac
 541 aagtgc当地 tcttacacca ggagaggatg gatgtttcgaa aactctatct tcactggcac
 601 accgtcgcca aagagacatg cagtgagaag agtaccaact tgcattgacta cggcatgttg
 661 ctggcctgca gaattgacaa gttccgaggg gtagagttt tgtagttgccc actggctgaa
 721 gaaagtgaca atgtggattc tgctgatgcg gaggaggatg actcggatgt ctggtgggc
 781 ggagcagaca cagactatgc agatgggatg gaagacaaag tagtagaaatg agcagaggag
 841 gaagaagtgg ctgaggtgga agaagaagaa gccgatgatg acgaggacga tgaggatggt
 901 gatgaggtag aggaagaggc tgaggaaccc tacgaagaag ccacagaag aaccaccac
 961 attgccacca ccaccaccac caccacagag tctgtgaag aggtggttcg agttcctaca
 1021 acagcagacca gtacccctga tgccgttgc aagtatctcg agacacctgg gtagatgaaat
 1081 gaacatgccc atttccagaa agccaaagag aggcttgagg ccaagcaccg agagagaatg
 1141 tcccaggtca tgagagaatg ggaagaggca gaacgtcaag caaagaactt gcciaaagct
 1201 gataagaagg cagttatcca gcatttccag gagaaagtgg aatctttgga acaggaagca
 1261 gccaacgaga gacagcagct ggtggagaca cacatggca gagtggaaacg catgtcaat
 1321 gaccggccgccc gcctggccct ggagaactac atcaccgctc tgcaaggctgt tcctcctcgg
 1381 cctcgtcacg tggtaatata gctaaagaag tatgtcccg cagaacagaa ggacagacag
 1441 cacaccctaa agcatttcga gcatgtgcgc atggtgatc ccaagaaagc cgctcagatc
 1501 cggccccagg ttatgacaca cctccgtgtt atttatgagc gcatgaatca gtctctctcc
 1561 ctgctctaca acgtgcctgc agtggccgag gagattcagg atgaagttga tgagctgctt
 1621 cagaaagagc aaaactattc agatgacgtc ttggcaaca tgattatgtga accaaggatc
 1681 agttacggaa acgatgctt catgccatct ttgaccgaaa cgaaaaccac cgtggagctc
 1741 cttccctgtga atggagagtt cagcctggac gatctccagc cgtggcattc ttttgggct
 1801 gactctgtc cagccaaacac agaaaacgaa gttgagcctg ttgatgccc ccctgctgcc
 1861 gaccgaggac tgaccactcg accaggttct gggttgacaa atatcaagac ggaggagatc
 1921 tctgaagtga agatggatgc agaattccga catgactcag gatatgaatg tcatcatcaa
 1981 aaattgggt tctttgcaga agatgtgggt tcaaacaag gtcaatcat tggactcatg
 2041 gtggcgggtt ttgtcatagc gacagtgtac gtcatcacct tggatgtgtc gaagaagaaa
 2101 cagtagacat ccattcatca tggtgtgggtt gaggttgacg ccgtgtcac cccagaggag
 2161 cggccacctgtt ccaagatgca gcagaacggc tacgaaaatc caacccatcaa gttctttag
 2221 cagatgcaga actagacccc cggccacagca gcctctgaag ttggacagca aaaccattgc
 2281 ttcaactaccc atcgggtgtcc atttatagaa taatgtggga agaaaacaaac ccgtttatg
 2341 atttactcat tatgccttt tgacagctgt gctgtacac aagtagatgc ctgaacttga
 2401 attaatccac acatcagtaa tgtattctat ctctcttac attttggctt ctataactaca
 2461 ttattaatgg gttttgtgtt ctgtaaagaa tttagctgtt tcaaactatgt gcatgaatag
 2521 attctctcctt gattattttt cacatggccc cttagccagt tggatattat tcttgggtt
 2581 tgtgacccaa ttaagtcccta ctttacatat gctttaagaa tcgatggggg atgcttcatg
 2641 tgaacgtggg agttcagctg cttctcttgc ctaagtattc ctttcctgtat cactatgcat
 2701 tttaaagttt aacatttttta agtatttcag atgcittaga gagatttttt ttccatgact
 2761 gcattttact gtacagattt ctgcttctgc tatatttgc atatagaaat taagaggata

FIG.18A

31/76

2821 cacacgtttg tttcttcgtg cctgttttat gtgcacacat taggcattga gacttcaagc
2881 ttttcttttt ttgtccacgt atctttgggt ctttgataaa gaaaagaatc cctgttcatt
2941 gtaagcactt ttacggggcg ggtggggagg ggtgctctgc tggtcttcaa ttaccaagaa
3001 ttctccaaaa caatttctg caggatgatt gtacagaatc attgctttag acatgatcgc
3061 tttctacact gtattacata aataaaattaa ataaaataac cccgggcaag actttcttt
3121 gaaggatgac tacagacatt aaataatcga agtaatttg ggtggggaga agaggcagat
3181 tcaattttct ttaaccagtc tgaagttca tttatgatac aaaagaagat gaaaatggaa
3241 gtggcaatat aaggggatga ggaaggcatg cctggacaaa cccttcttt aagatgtgtc
3301 ttcaatttgt ataaaaatggt gttttcatgt aaataaaatac attcttggag gagc

FIG. 18B

32/76

(SEQ ID NO: 18)

MLPGLALLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMH
MNVQNGKWDSDPSGTKTCIDTKEGILQYCQEYPELQITNVVEANQPVTIQNWCKRGR
KQCKTHPHFVIPYRCLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSE
KSTNLHDYGMLLPCGIDKFRGVEVCCPLAEESDNVDSDADAEEEDDSDVWWGGADTDYA
DGSEDKVVEAEEEEVAEVEEEEADDDEDGDEVEEEAEPYEEATERTTSIATTT
TTTTEVESVEVVRVPTTAASTPDADVCKYLETPGDENEHAHFQKAKERLEAKHRERMSQV
MREWEEEAERQAKNLPKADKKAVIQHFQEKFVESLEQEAANERQQLVETHMARVEAMLND
RRRLALENYITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKFEHVRMVDPKKAQ
IRSQVMTHLRVIYERMNQSLSLYNPVAEEIQDEVDELLQKEQNYSDDVLANMISE
PRISYGNDALMPSLTETKTTVELLPVNGEFSLDDLQPWHSGADSVPANTENEVEPVD
ARPAADRGLTTRPGSGLTNIKTEEISEVKMDAÉFRHDSGYEVHHQKLVFFAEDVGSNK
GAIIGLMVGGVVIATVIVITLVMKKQYTSIHGVVEVDAAVTPEERHLSKMQQNGY
ENPTYKFFEQMQN

FIG.18C

33/76

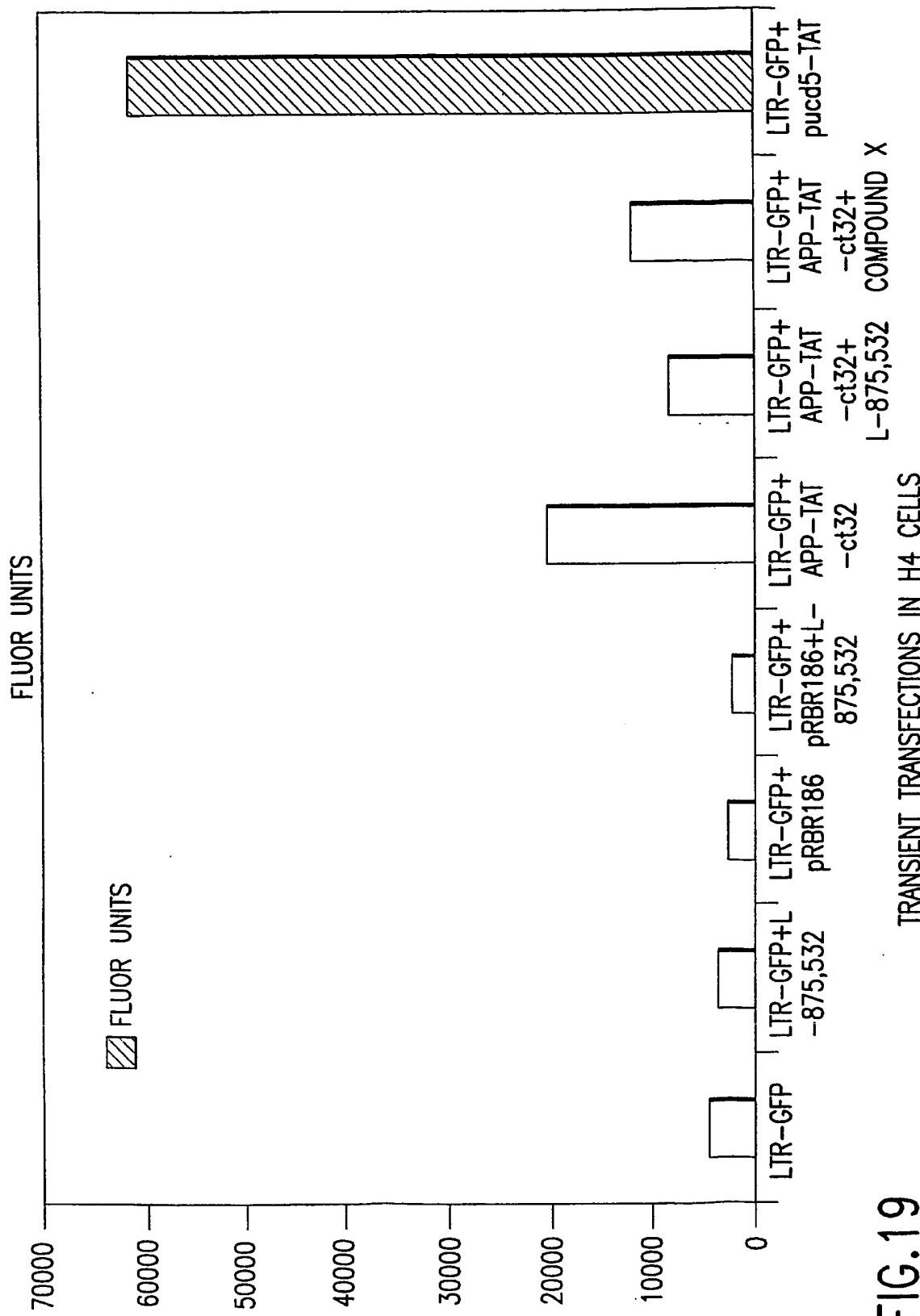


FIG. 19

34/76

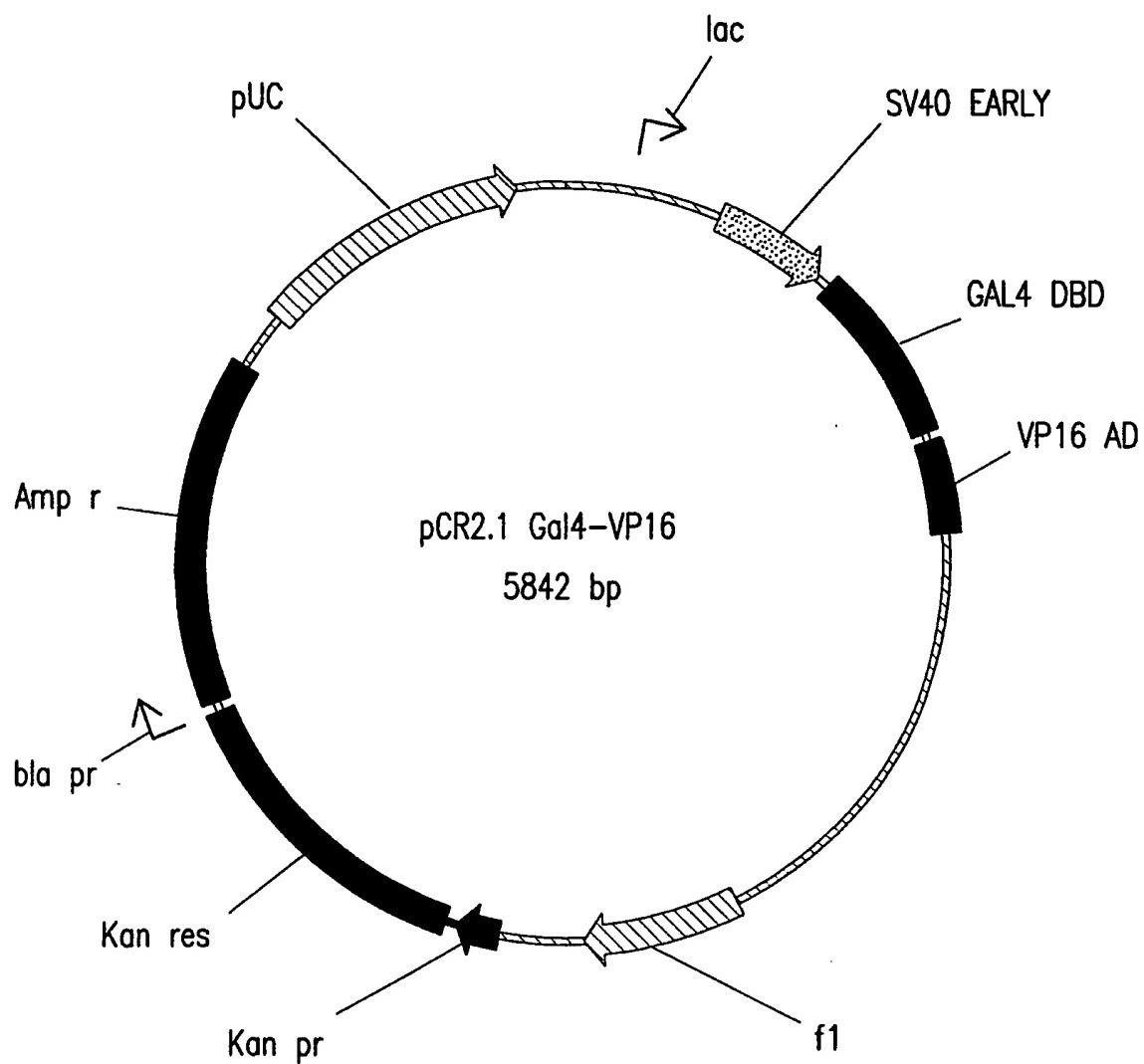


FIG.20

35/76

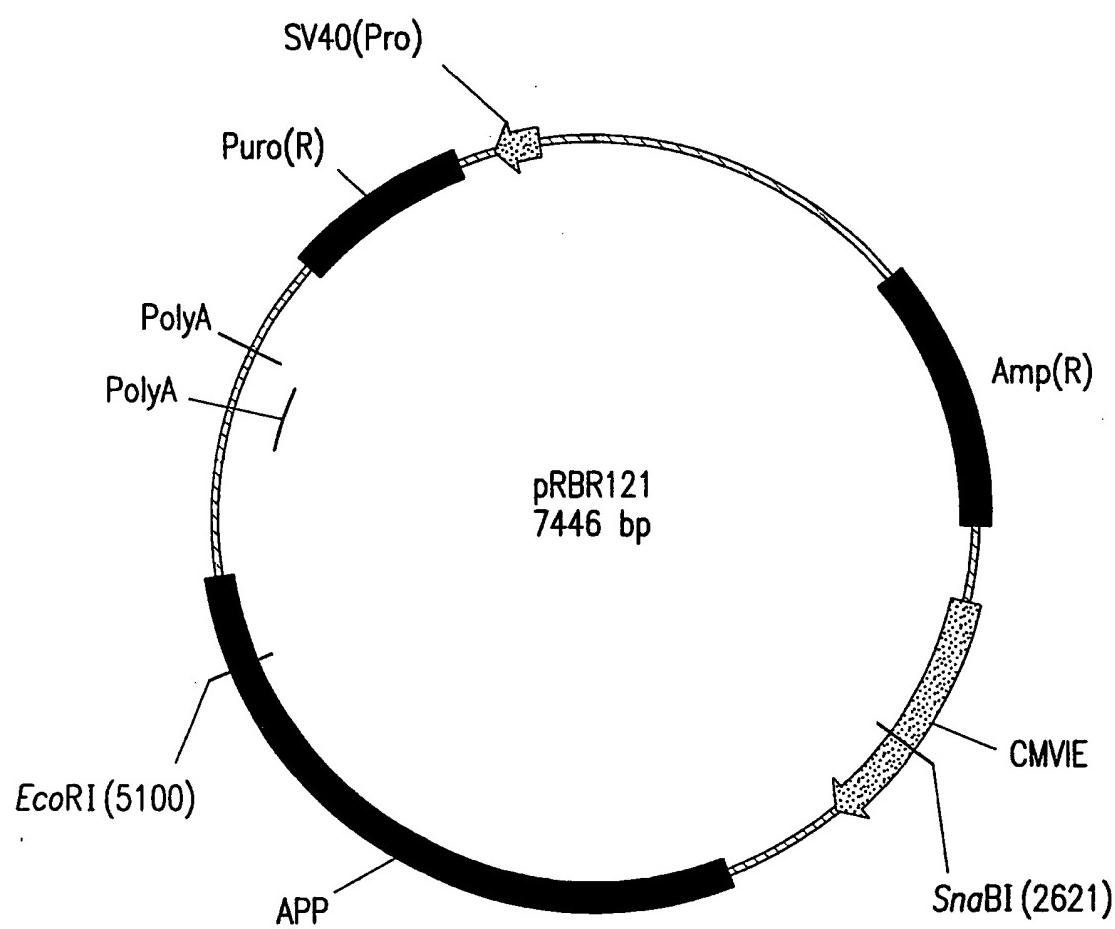


FIG.21A

36/76

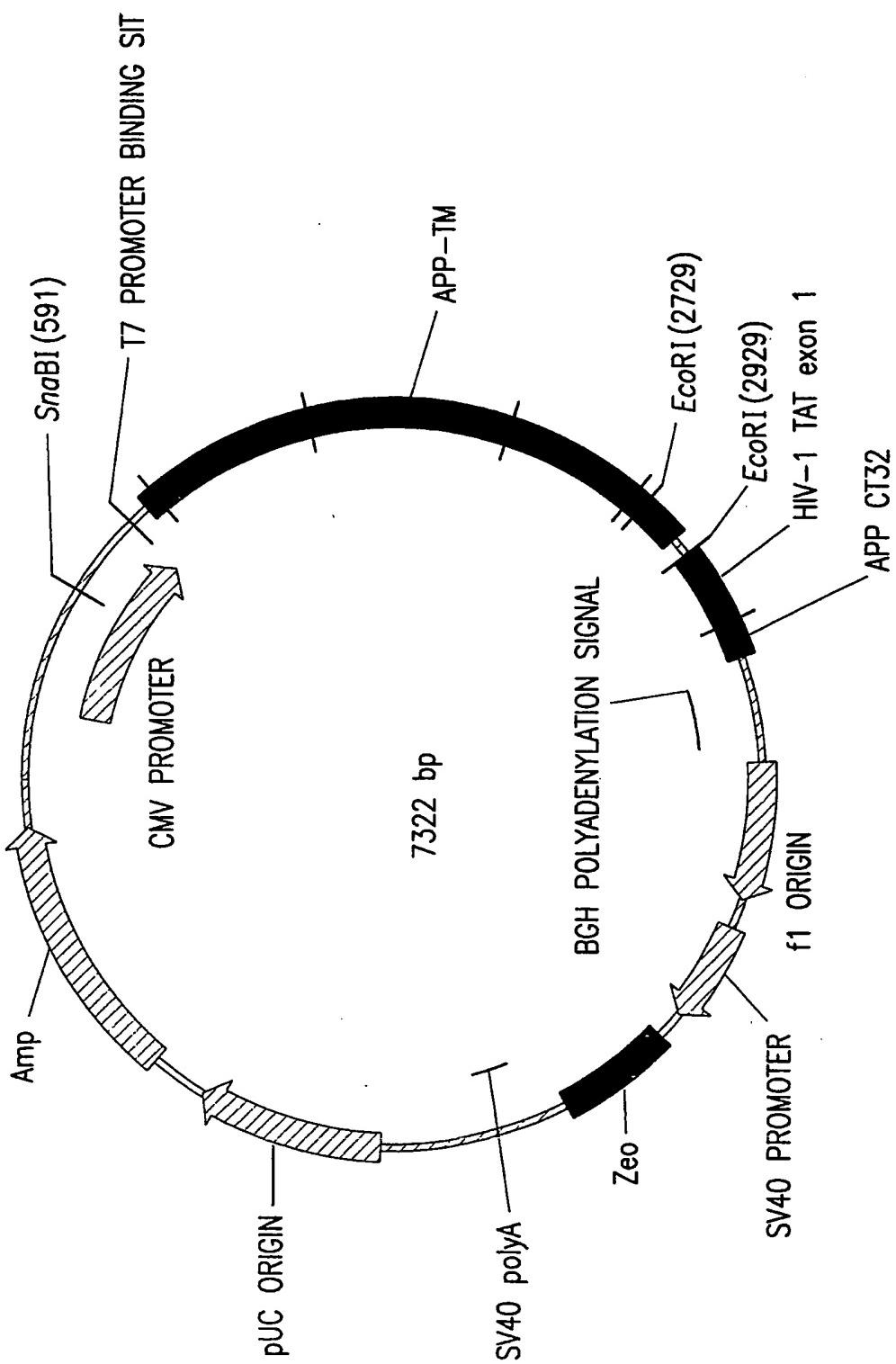


FIG. 21B

37/76

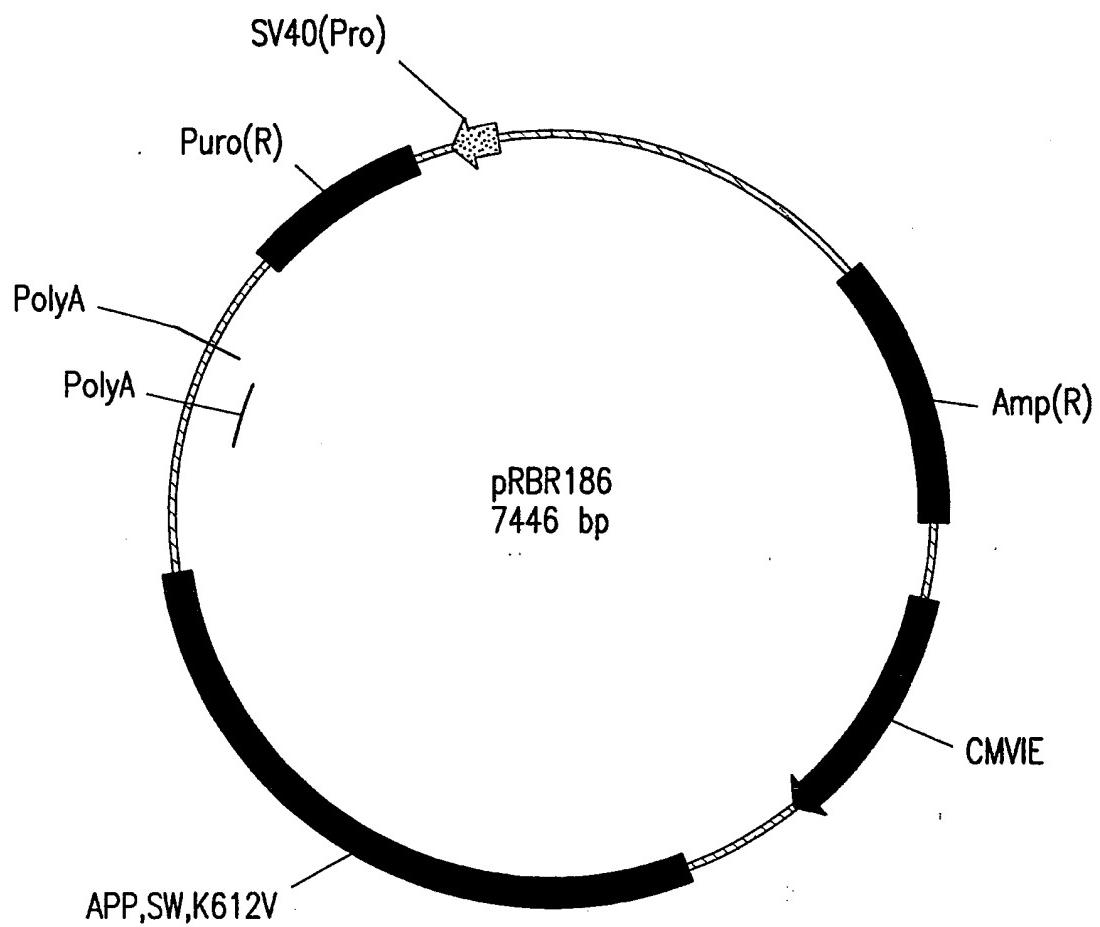
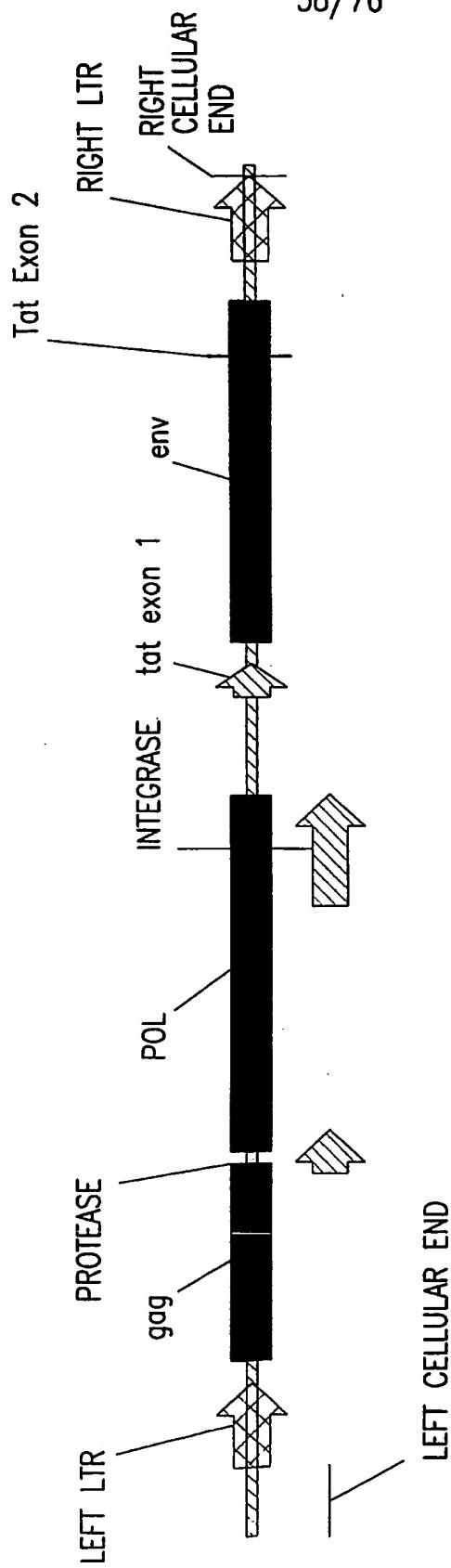


FIG.22A

38/76



10317 bp

FIG. 22B

39/76

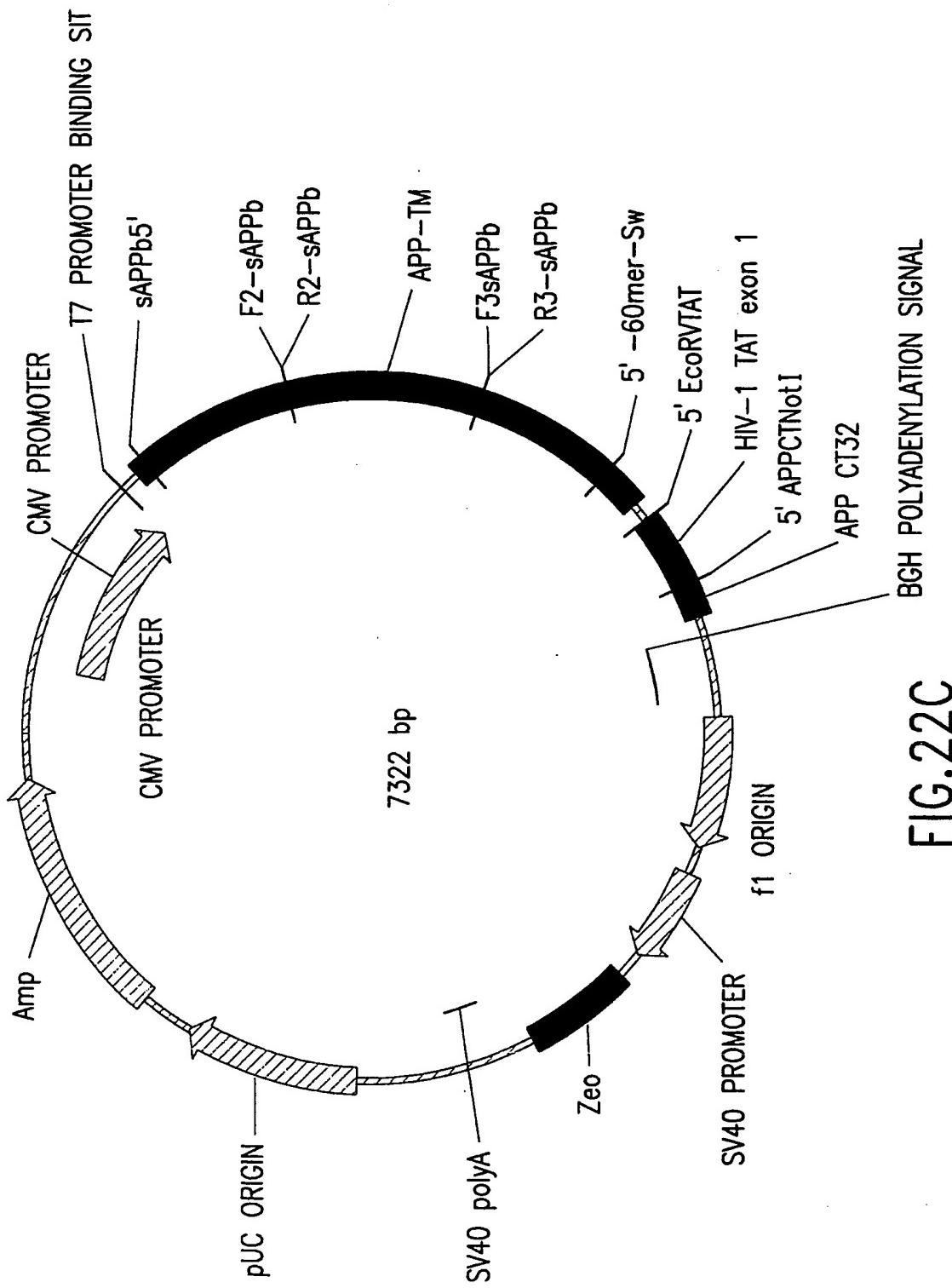


FIG. 22C

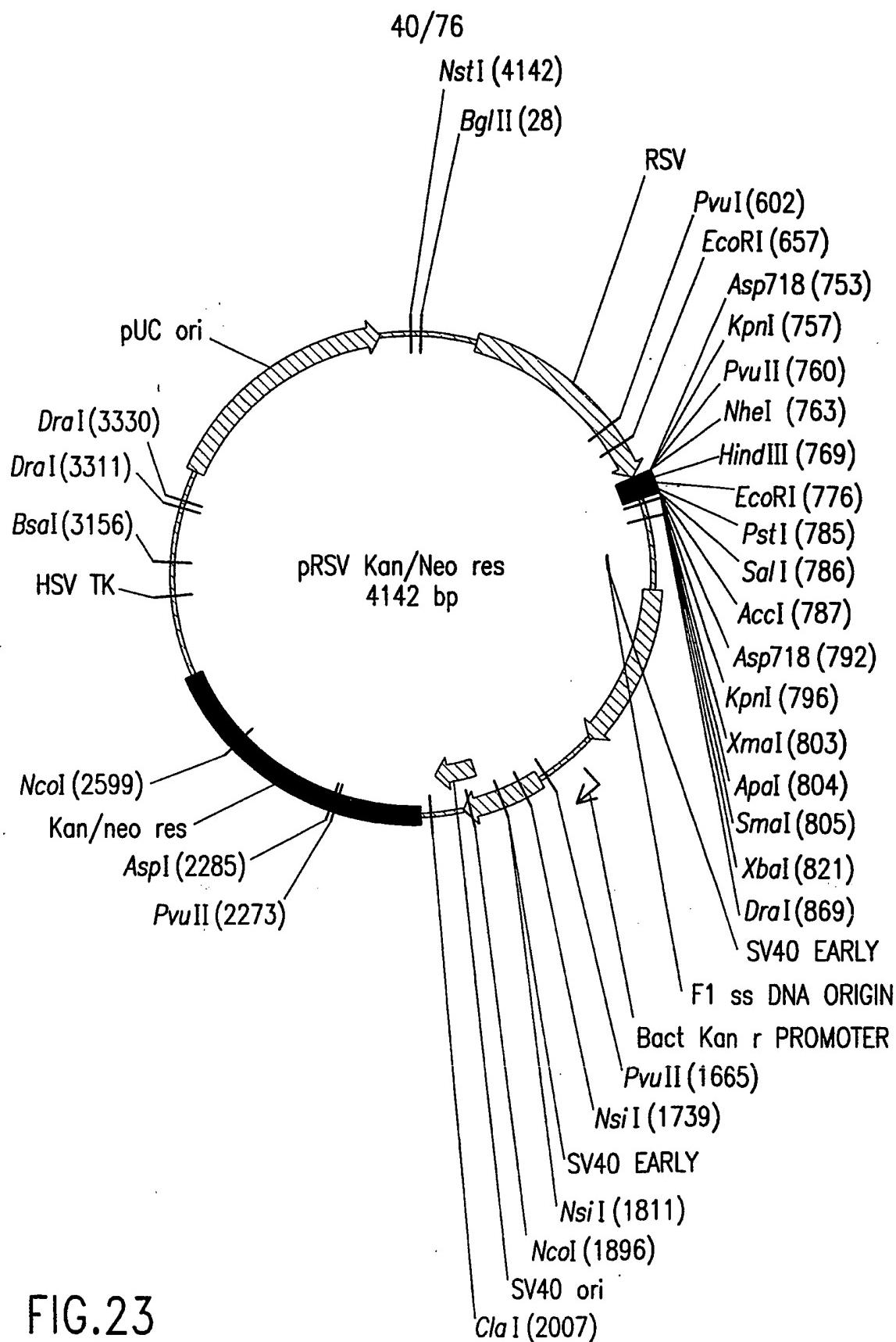


FIG.23

41/76

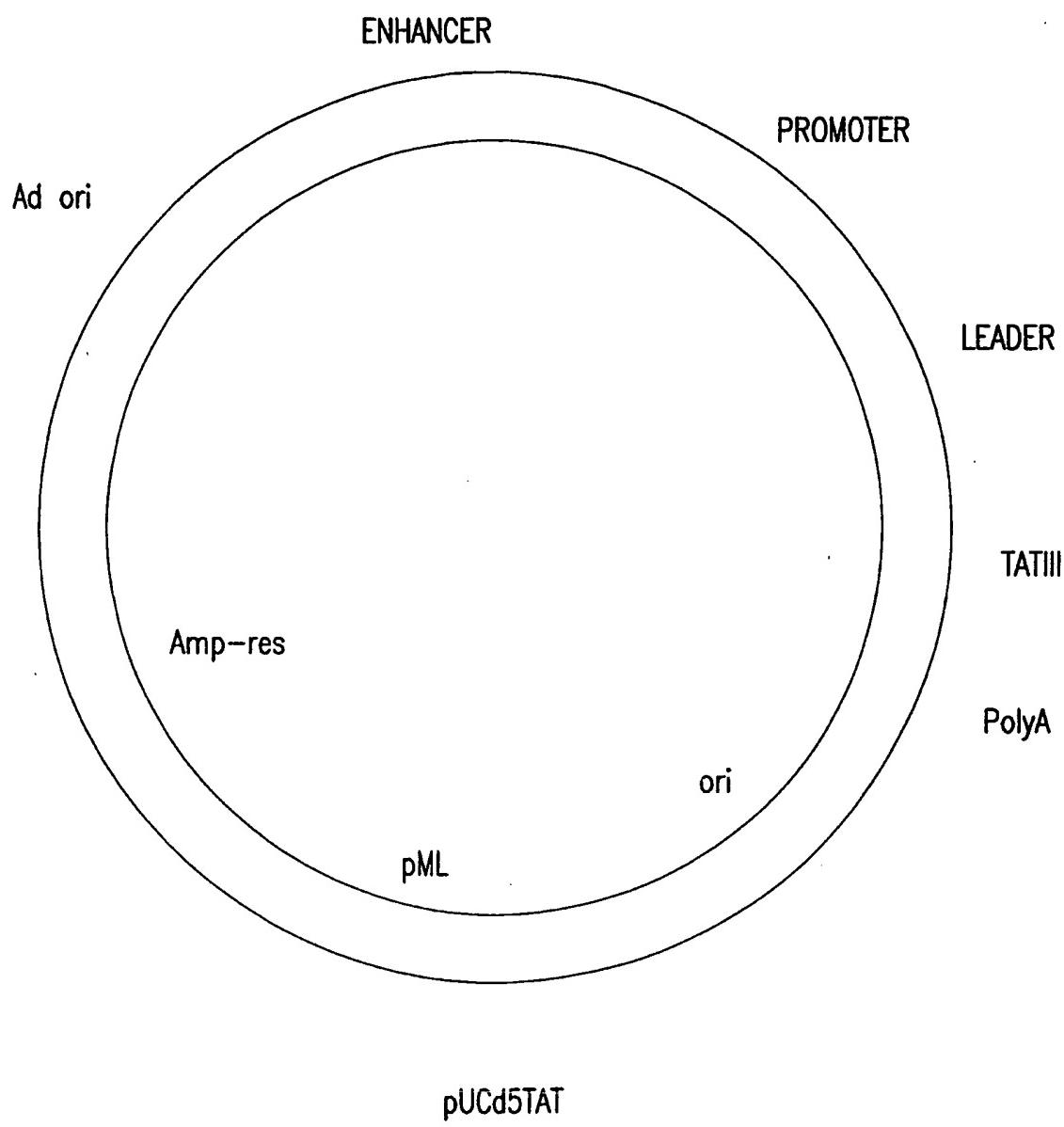


FIG.24

42/76

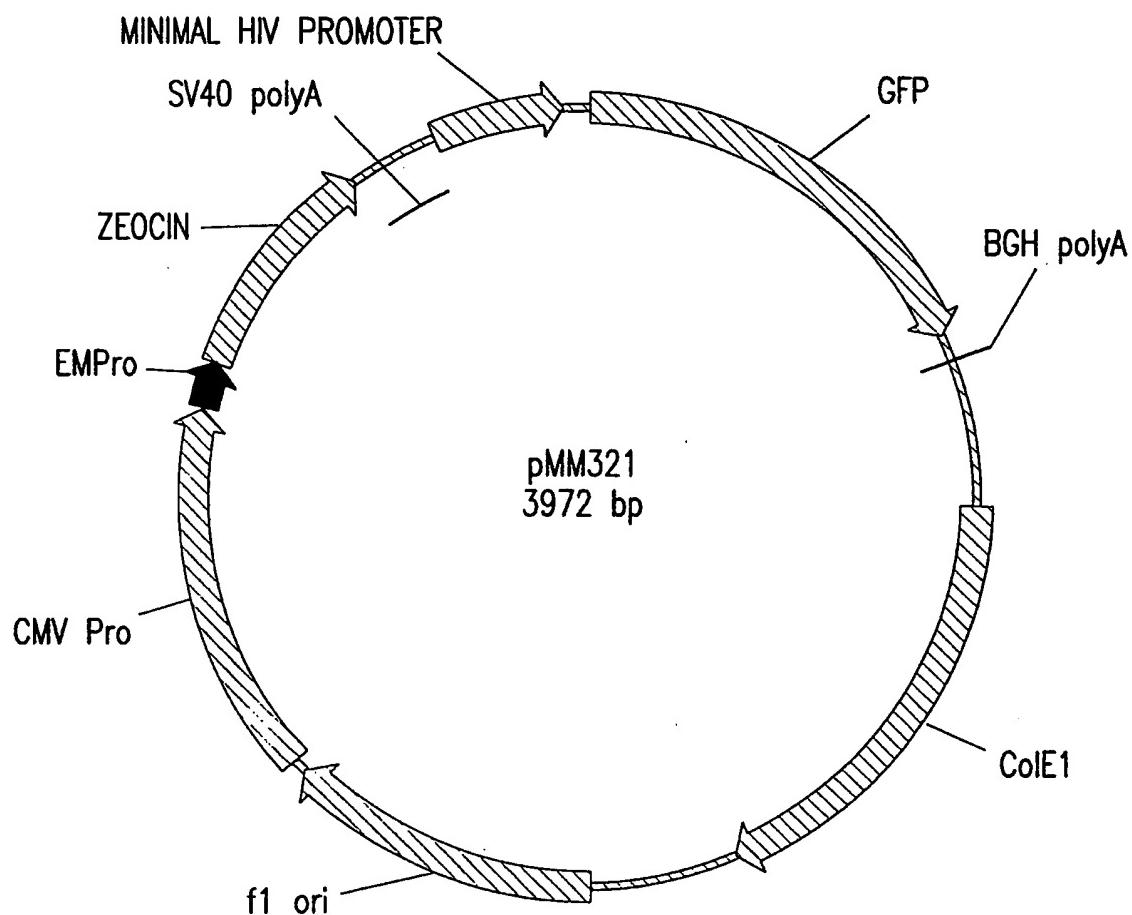


FIG.25A

43/76

(SEQ ID NO: 19 AND 20)

1	ATGGTGAGCA	AGGGCGAGGA	GCTGTTCAC	GGGGTGGTGC	CCATCCTGGT
	TACCACTCGT	TCCCGCTCCT	CGACAAGTGG	CCCCACCAACG	GGTAGGACCA
51	CGAGCTGGAC	GGCGACGTA	ACGCCACAA	GTTCAGCGTG	TCCGGCGAGG
	GCTCGACCTG	CCGCTGCATT	TGCGGGTGT	CAAGTCGCAC	AGGCCGCTCC
101	GGCAGGGCGA	TGCCACCTAC	GGCAAGCTGA	CCCTGAAGTT	CATCTGCACC
	CGCTCCCGCT	ACGGTGATG	CCGTTGACT	GGGACTTCAA	GTAGACGTGG
151	ACCGGCAAGC	TGCCCCTGCC	CTGGCCACC	CTCGTGACCA	CCTTCACCTA
	TGGCCGTTCG	ACGGGCACGG	GACCGGGTGG	GAGCACTGTT	GGAAAGTGGAT
201	CGGCGTGCAG	TGCTTCGCC	GCTACCCCGA	CCACATGAAG	CAGCACGACT
	GCCGCACGTC	ACGAAGCGGG	CGATGGGGCT	GGTGTACTTC	GTCGTGCTGA
251	TCTTCAAGTC	CGCCATGCC	GAAGGCTACG	TCCAGGAGCG	CACCATCTTC
	AGAAGTTCA	CGGGTACCGG	CTTCCGATGC	AGGTCTCTCG	GTGGTAGAAG
301	TTCAGGACG	ACGGCAACTA	CAAGACCCGC	GCCGAGGTTGA	AGTTCGAGGG
	AAGTTCTCG	TGCCGTTGAT	GTTCTGGGCG	CGGCTCCACT	TCAAGCTCCC
351	CGAACACCTG	GTGAAACCGCA	TCGAGCTGAA	GGGCATCGAC	TTCAGGAGG
	GCTGTGGGAC	CACTTGGCGT	AGCTGACTT	CCCGTAGCTG	AAGTTCTCC
401	ACGGCAACAT	CCTGGGCAC	AAGCTGGAGT	ACAACATACAA	CAGCCACAAG
	TGCCGTTGTA	GGACCCCGTG	TTGACCTCA	TGTTGATGTT	GTCGGTGTTC
451	GTCTATATCA	CCGCCGACAA	GCAGAAGAAC	GGCATCAAGG	TGAACATTCA
	CAGATATAGT	GGCGGCTGTT	CGTCTTCTTG	CCGTAGTTCC	ACTTGAAGTT
501	GACCCGCCAC	AACATCGAGG	ACGGCAGCGT	GCAGCTCGCC	GACCACTACC
	CTGGGCGGTG	TTGTAGCTC	TGCCGTCGA	CGTCGAGCGG	CTGGTGTATGG
551	AGCAGAACAC	CCCCATCGGC	GACGGCCCG	TGCTGCTGCC	CGACAACCCAC
	TCGTCTTGTG	GGGGTAGCCG	CTGCCGGGGC	ACGACGACGG	GCTGTTGGTG
601	TACCTGAGCA	CCCAGTCCG	CCTGAGCAAA	GACCCCAACG	AGAAGCGCGA
	ATGGACTCGT	GGGTCAAGCG	GGACTCGTT	CTGGGGTTGC	TCTTCGCGCT
651	TCACATGGTC	CTGCTGGAGT	TGCTGACCGC	CGCCGGGATC	ACTCTCGGCA
	AGTGTACCAAG	GACGACCTCA	AGCACTGGCG	GGGGCCCTAG	TGAGAGCCGT
701	TGGACGAGCT	GTACAAGTAA	CTCGAGTCTA	GAGGGCCCGT	TTAAACCCCGC
	ACCTGTCGA	CATGTTTATT	GAGCTCAGAT	CTCCCGGGCA	AATTGGGCG
751	TGATCAGCCT	CGACTGTGCC	TTCTAGTTGC	CAGCCATCTG	TTGTTTGGCC
	ACTAGTCGGA	GCTGACACGG	AAGATCAACG	GTCGGTAGAC	AACAAACGGG
801	CTCCCCCGTG	CCTTCCTTGA	CCCTGGAAGG	TGCCACTCCC	ACTGTCCTTT
	GAGGGGGGAC	GGAAAGGAAC	GGGACCTTCC	ACGGTGAGGG	TGACAGGAAA
851	CCTAATAAAA	TGAGGAAATT	GCATCGCATT	GTCTGAGTAG	GTGTCATTCT
	GGATTATTTT	ACTCTTTAA	CGTAGCGTAA	CAGACTCATC	CACAGTAAGA
901	ATTCTGGGGG	GTGGGGTGGG	GCAGGACAGC	AAGGGGGAGG	ATTGGGAAGA
	TAAGACCCCC	CACCCACCC	CGTCTGTGCG	TTCCCCCTCC	TAACCTTCT
951	CAATAGCAGG	CATGCTGGG	ATGCGGTGGG	CTCTATGGCT	TCTGAGGCGG
	GTTATCGTCC	GTACGACCCC	TACGCCACCC	GAGATACCGA	AGACTCCGCC
1001	AAAGAACCCAG	CATGTGAGCA	AAAGGCCAGC	AAAGGCCAG	GAACCGTAAA
	TTCTTGGTC	GTACACTCGT	TTTCGGTGC	TTTCGGTTC	CTTGGCATT
1051	AAGGCCGCGT	TGCTGGCGT	TTTCCATAGG	CTCCGCC	CTGACGAGCA
	TTCCGGCGCA	ACGACCGCAA	AAAGGTATCC	GAGGCGGGGG	GACTGCTCGT
1101	TCACAAAAAT	CGACGCTCAA	GTCAGAGGTG	GCGAAACCCG	ACAGGACTAT
	AGTGTTTTTA	GCTGCGAGTT	CAGTCTCCAC	CGCTTGGC	TGTCTGATA
1151	AAAGATACCA	GGCGTTTCCC	CCTGGAAGCT	CCCTCGTGC	CTCTCTGTT
	TTTCTATGGT	CCGCAAAGGG	GGACCTTCA	GGGAGCAGC	GAGAGGACAA
1201	CCGACCCCTGC	CGCTTACCGG	ATACCTGTCC	GCCTTCTCC	CTTCGGGAAG
	GGCTGGGACG	GGCAATGGC	TATGGACAGG	CGGAAAAGAGG	GAAGCCCTTC
1251	CGTGGCGCTT	TCTCATAGCT	CACGCTGTAG	GTATCTCA	TCGGTGTAGG
	GCACCCCGAA	AGAGTATCGA	GTGCGACATC	CATAGAGTCA	AGCCACATCC
1301	TCGTTCGCTC	CAAGCTGGC	TGTGTGCACG	AACCCCCCGT	TCAGCCCGAC
	AGCAAGCGAG	GTTCGACCGC	ACACACGTGC	TTGGGGGGCA	AGTCGGGCTG
1351	CGCTGCGCT	TATCCGGTAA	CTATCGTCTT	GAGTCCAACC	CGGTAAGACA
	CCGACGCGGA	ATAGGCATT	GATAGCAGAA	CTCAGGTTGG	GCCATTCTGT
1401	CGACTTATCG	CCACTGGCAG	CAGCCACTGG	TAACAGGATT	AGCAGAGCGA
	GCTGAATAGC	GGTGACCGTC	GTCGGTGACC	ATTGTCTTAA	TGTCCTCGCT

FIG.25B

44/76

1451	GGTATGTAGG	CGGTGCTACA	GAGTTCTTGA	AGTGGTGGCC	TAAC TACGGC
	CCATACATCC	GCCACGATGT	CTCAAGAACT	TCAACCACCGG	ATTGATGCCG
1501	TACACTAGAA	GAACAGTATT	TGGTATCTGC	GCTCTGCTGA	AGCCAGTTAC
	ATGTGATCTT	CTTGTATAA	ACCATAGACG	CGAGACGACT	TCGGTCAATG
1551	CTTCGGAAAA	AGAGTTGGTA	GCTCTTGATC	CGGCAAACAA	ACCACCGCTG
	GAAGCCTTT	TCTCAACCAT	CGAGAACTAG	GCCGTTTGT	TGGTGGCGAC
1601	GTAGCGGTGG	TTTTTTTGT	TGCAAGCAGC	AGATTACGCG	CAGAAAAAAA
	CATCGCCACC	AAAAAAACAA	ACGTTCGTCG	TCTAATGCGC	GTC
1651	GGATCTCAAG	AAGATCCTT	GATTTTCT	ACGGGGTCTG	ACGCTCAGTG
	CCTAGAGTTC	TTCTAGGAAA	CTAGAAAAGA	TGCCCCAGAC	TGCCAGTCAC
1701	GAACGAAAAC	TCACGTTAAG	GGATTTGGT	CATGACATTA	ACCTATAAAA
	CTTGCTTTTG	AGTGCAATT	CCTAAACCA	GTACTGTAA	TGGATATTTT
1751	ATAGGCGTAT	CACGAGGCC	TTTCGTC	CGCGTTT	TGATGACGGT
	TATCCGCATA	GTGCTCCGGG	AAAGCAGAGC	GCGCAAAGCC	ACTACTGCCA
1801	GAAAACCTCT	GACACATGCA	GCTCCCGGAG	ACGGTCACAG	CTTGTCTGTA
	CTTGGGAGA	CTGTTGTA	CGAGGGCCTC	TGCCAGTGTC	GAACAGACAT
1851	AGCGGAATGCC	GGGAGCAGAC	AAGCCGTCA	GGGCGCGTCA	GCGGGTGTG
	TCGCCTACGG	CCCTCGTCTG	TTCGGCAGT	CCCGCGCAGT	CGCCCACAAAC
1901	GCGGGTGTGCG	GGGCTGGCTT	AACTATGCGG	CATCAGAGCA	GATTGTA
	CGCCCACAGC	CCCGACCGAA	TTGATACGCC	GTAGTCTCGT	CTAACATGAC
1951	AGAGTGCACC	ATATGCGGT	TGAAATACCG	CACAGATGCG	TAAGGAGAAA
	TCTCACGTGG	TATACGCCAC	ACTTTATGCC	GTGTCTACGC	ATTCCTCTTT
2001	ATACCGCATH	AGGACGCGCC	CTGTAGCGGC	GCATTAAGCG	CGGCAGGTGT
	TATGGCGTAG	TCCTGCGCGG	GACATCGCCG	CGTAATT	GCCGCCACAA
2051	GGTGGTTACG	CGCAGCGTGA	CGCCTACACT	TGCCAGCGCC	CTAGCGCCCG
	CCACCAATGC	CGCTCGCACT	GGCGATGTGA	ACGGTCGCGG	GATCGCGGGC
2101	CTCTTTTCG	TTTCTTCCCT	TCCTTTCTCG	CCACGTT	CGGCTTTCCC
	GAGGAAAGCG	AAAGAAGGGA	AGGAAAGAGC	GGTGAAGCG	GCGAAAGGG
2151	CGTCAAGCTC	TAATCGGGG	GCTCCCTTA	GGGTTCCGAT	TTAGTGTCTT
	GCAGTTGAG	ATTTAGCCCC	CGAGGGAAAT	CCCAAGGCTA	AATCACGAAA
2201	ACGGCACCTC	GACCCCAAAA	AACTTGATTA	GGGTGATGGT	TCACGTAGTG
	TGCCGTGGAG	CTGGGGTTT	TTGAACTAAT	CCCACTACCA	AGTGCATCAC
2251	GGCATCGCC	CTGATAGACG	GTTTTCGCC	CTTGTACGTT	GGAGTCCACG
	CCGGTAGCGG	GACTATCTG	CAAAAGCGG	GAAACTGCAA	CCTCAGGTGC
2301	TTCTTTAATA	GTGGACTCTT	GTTCCA	GGAACAACAC	TCAACCCAT
	AAGAAATTAT	CACCTGAGAA	CAAGGTTGA	CCTTGTG	AGTTGGGATA
2351	CTCGGTCTAT	TCTTTGATT	TATAAGGGAT	TTGCGCATT	TCGGCCTATT
	GAGCCAGATA	AGAAAACAA	ATATCCCTA	AAACGGCTAA	AGCCGGATAA
2401	GGTTAAAAAA	TGAGCTGATT	TAACAAAAT	TTAACGCGAA	TTTAACAAA
	CCAA	ACTCGACTAA	ATTGTTTTA	AATTGCGCTT	AAAATTGTT
2451	ATATTAACGC	TTACAATTTC	CATTGCCAT	TCAGGCTGAA	CTAGATCTAG
	TATAATTGCG	AATGTTAAAG	GTAAGCGGT	AGTCCGACTT	GATCTAGATC
2501	AGTCGTTAC	ATAACTTACG	GTAATGGCC	CGGCTGGCTG	ACCGCCCCAAC
	TCAGGCAATG	TATTGAATGC	CATTACCGG	GCGGACCGAC	TGGCGGGTTG
2551	GACCCCGGCC	CATTGACGTC	AATAATGACG	TATGTTCCCA	TAGTAACGCC
	CTGGGGCGG	GTAACTGCGAG	TTATTACTGC	ATACAAGGGT	ATCATTGCGG
2601	AATAGGGACT	TTCCATTGAC	GTCAATGGGT	GGAGTATTTA	CGGTAAACTG
	TTATCCCTGA	AAGGTA	CAGTTACCA	CCTCATAAAT	GCCATTGAC
2651	CCCACTGGC	AGTACATCAA	GTGTATCATA	TGCCAAGTAC	GCCCCCTATT
	GGGTGAACCG	TCATGTAGTT	CACATAGTAT	ACGGTTCATG	GGGGGGATAA
2701	GACGTCATG	ACGGTAAATG	GCCCCGCTGG	CATTATGCC	AGTACATGAC
	CTGAGTAC	TGCCATTAC	CGGGCGGACC	GTAATACGGG	TCATGTACTG
2751	CTTATGGGAC	TTTCTACTT	GGCAGTACAT	CTACGTATTA	GTCATCGCTA
	GAATACCCCTG	AAAGGATGAA	CCGTCA	GATGCATAAAT	CAGTAGCGAT
2801	TTACCATGGT	GATGCGGTT	TGGCAGTACA	TCAATGGGCG	TGGATAGCGG
	AATGGTACCA	CTACGCCAAA	ACCGTCATGT	AGTACCCGC	ACCTATGCC
2851	TTTGA	GGGATTTC	AAGTCTCCAC	CCCATTGACG	TCAATGGGAG
	AAACTGAGTG	CCCCTAAAGG	TTCAGAGGTG	GGGTA	AGTACCCCTC

FIG.25C

45/76

2901 TTTGTTTG G CACCAAAATC AACGGGACTT TCCAAAATGT CGTAACAAC
 2951 AAAACAAAACC GTGGTTTAG TTGCCCTGAA AGGTTTTACA GCATTGTTGA
 2951 CGGCCCTTATT GACGCAAATG GGCGGTAGGC GTGTACGGTG GGAGGTCTAT
 3001 GGCGGGGTAA CTGCGTTTAC CGGCATCCG CACATGCCAC CCTCCAGATA
 3001 ATAAGCAGAG CTCGTTAGT GAACCGTCAG ATCGCCTGGA GACGCCATCC
 3051 TATTCTGTC GAGCAAATCA CTTGGCAGTC TAGCGGACCT CTGCGGTAGG
 3051 ACGCTGTTT GACCTCCATA GAAGACACCG GGACCGATCC AGCCTCCGCG
 3101 TGCAGACAAAA CTGGAGGTAT CTTCTGTGGC CCTGGCTAGG TC GGAGGCGC
 3101 GCCGGGAACG GTGCATTGGA ACGGACCGTG TTGACAATTAA ATCATCGGCA
 3151 CGGCCCTTGC CACGTAACCT TGCCCTGGCAC AACTGTAAAT TAGTAGCCGT
 3151 TAGTATATCG GCATAGTATAA ATACGACAAG GTGAGGAACAAACCATGGC
 3201 ATCATATAGC CGTATCATAT TATGCTGTTC CACTCCTTGA TTGGTACCG
 3201 CAAGTTGACC AGTGCCGTT TCACGGCAAG GCCACGAGTG GCGCGCGCTG CAGCGGCCTC
 3251 GTTCAACTGG CGGTGAGTT CTGGACCGAC CGGCTCGGGT TCTCCCAGGG CTTCTGGAG
 3251 GCCAGCTCAA GACCTGGCTG GCGGAGGCCA AGAGGGCCCT GAAGCACCTC
 3301 GACGACTTCG CCGGTGTGGT CCGGGACGAC GTGACCCCTGT TCATCAGCGC
 3301 CTGCTGAAGC GGCCACACCA GGCCCTGCTG CACTGGGACA AGTAGTCGCG
 3351 3351 GGTCCAGGAC CAGGTGGTGC CGGACAAACAC CCTGGCCTGG GTGTGGGTGC
 CCAGGTCTG GTCCACCAAG GCCTGGTGTG GGACCGGACCC CACACCCACG
 3401 CGGGCCTGGA CGAGCTGTAC GCGAGGTGGT CGGAGGTCTG GTCCACGAAAC
 CGCCGGACCT GCTCGACATG CGGCTCACCA GCTCCAGCA CAGGTGCTTG
 3451 3451 TTCCGGGACG CCTCCGGGCC GGCCATGACC GAGATCGGCG AGCAGCCGTG
 AAGGGCCTGC GGAGGCCCCGG CCGGTACTGG CTCTAGCCGC TCGTCGGCAC
 3501 3501 GGGGGGGAG TTGCCCCCTGC GCGACCCGGCG CGCACAATGCG GTGCACTTCG
 CCCCGCCCTC AAGCGGGACG CGCTGGGGCGG GCCGTTGACG CACGTGAAGC
 3551 3551 TGGCCGAGGA GCAGGACTGA CACTCGACCT CGAAACTTGT TTATTGAGC
 ACCGGCTCT CGTCCCTGACT GTGAGCTGGA GCTTGAACA AATAACGTCG
 3601 3601 TTATAATGGT TACAAATAAA GCAATAGCAT CACAAATTTC ACAAAATAAG
 AATATTACCA ATGTTTATTT CGTTATCGTA GTGTTTAAAG TGTTTATTC
 3651 3651 CATTTTTTTC ACTGCATTCT AGTTGTGGTT TGTC CAAACT CATCAATGTA
 GTAAAAAAAG TGACGTAAGA TCAACACCAA ACAGGTTTGA GTAGTTACAT
 3701 3701 TCTTATCATG TCTGGATCGA TACTTCAAGA ACTGCTGACA TCGAGCTTGC
 AGAATAGTAC AGACCTAGCT ATGAAGTTCT TGACGACTGT AGCTCGAACG
 3751 3751 TACAAGGGAC TTTCCGCTGG GGACTTTCCA GGGAGGCGTG GCCTGGCGG
 ATGTTCCCTG AAAGGCACCC CCTGAAAGGT CCCTCCGCAC CGGACCCGCC
 3801 3801 GACTGGGGAG TGGCGAGGCC TCAGATCCTG CATATAAGCA GCTGCTTTT
 CTGACCCCTC ACCGCTCGGG AGTCTAGGAC GTATATTCTG CGACGAAAAAA
 3851 3851 GCCTGTACTG GGTCTCTCTG GTTAGACCGAG ATCTGAGGCCT GGGAGCTCTC
 CGGACATGAC CCAGAGAGAC CAATCTGGTC TAGACTCGGA CCCTCGAGAG
 3901 3901 TGGCTAACTA GGGAAACCCAC TGCTTAAGCC TCAATAAGC TTGGTACCGA
 ACCGATTGAT CCCTTGGGTG ACGAATTCTGG AGTTATTTCG AACCATGGCT
 3951 3951 GCTCGGATCC GAATTGCGCA CC CGAGCCTAGG CTTAAGCGGT GG

FIG.25D

46/76

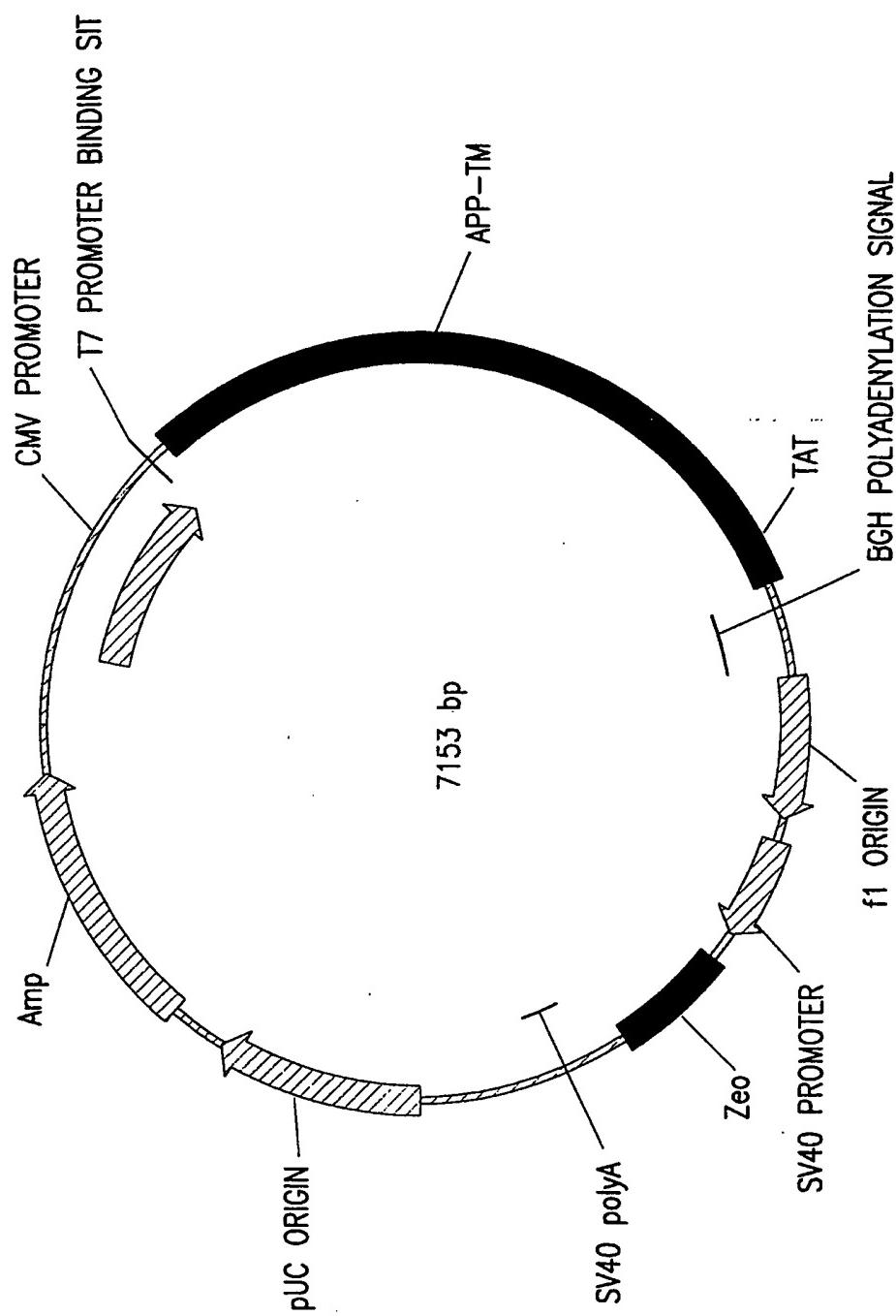


FIG. 26A

47/76

(SEQ ID NO: 21 AND 22)

1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC
 CTGCCTAGCC CTCTAGAGGG CTAGGGGATA CCAGCTGAGA GTCACTGTTAG
 51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT
 ACGAGACTAC GGCCTATCAA TTCGGTCATA GACGAGGGAC GAACACACAA
 101 GGAGGTCGCT GAGTAGTGC CGAGCAAAAT TTAAGCTACA ACAAGGCAAG
 CCTCCAGCGA CTCATCACCG GCTCGTTTA ATTCTGATGT TGTTCCGTT
 151 GCTTGACCGA CAATTGCACTG AAGAACATCG TTAGGGTTAG GCGTTTTGCG
 CGAACCTGGCT GTTAACGTC TTCTTAGACG AATCCCAATC CGCAAAACGC
 201 CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT GATTATTGAC
 GACGAAGCGC TACATGCCCG GTCTATATGC GCAACTGTAA CTAATAACTG
 251 TAGTTATTAAT TAGTAATCAA TTACGGGTC ATTAGTTCAT AGCCCATATA
 ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCAGGGTATAT
 301 TGAGGTTCCG CGTTACATCAA CTTACGGTAA ATGGCCGCC TGCTGACCG
 ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC
 351 CCCAACGACC CCCGCCATT GACGTCATAA ATGACGTATG TTCCCATAGT
 GGTTGCTGG GGGCGGGTAA CTGCACTTAT TACTGCATAC AAGGGTATCA
 401 AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAC TATTTACGGT
 TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTG ATAAATGCCA
 451 AACTGCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC
 TTTGACGGGT GAACCGTCAT GTAGTTACA TAGTATACGG TTCACTGCGG
 501 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA
 GGATAACTGC AGTTACTGCC ATTACCGGG CGGACCGTAA TACGGGTCA
 551 CATGACCTTA TGGGACTTTCT ACTCTGGCA GTACATCTAC GTATTAGTCA
 GTACTGGAAT ACCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT
 601 TCGCTATTAC CATGGTGTATG CGGTTTGGC AGTACATCAA TGGGCGTGG
 AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 651 TAGCGGTTTG ACTCACGGGG ATTTCAAGT CTCCACCCCA TTGACGTAA
 ATCGCCAAAC TGAGTCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT
 701 TGGGAGTTTG TTTTGGCACCA AAAATCAACG GGACTTTCCA AAATGTCGA
 ACCCTCAAAC AAAACCGTGG TTTAGTTGC CCTGAAAGGTT TTTACAGCAT
 751 ACAACTCCGC CCCATTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG
 TGTTGAGGCG GGGTAACTGC GTTACCCGC CATCCGCACA TGCCACCC
 801 GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA CTGCTTACTG
 CAGATATATT CGTCTCGAGA GACCGATTGA TCTCTGGGT GACGAATGAC
 851 GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
 CGAACATAGCTT TAATTATGCT GAGTGTATC CCTCTGGGTT CGACCGATCG
 901 GTTTAAACTT AAGCTTCCCC GCGCAGGGTC GCGATGCTGC CCGGTTGGC
 CAAATTGAA TTCAAGGGG CGCGTCCCAG CGCTACGACG GGCCAAACCG
 951 ACTGCTCCTG CTGGCCGCCT GGACGGCTCG GGCCTGGAG GTACCACTG
 TGACGAGGAC GACCGGGCGA CCTGCCGAGC CCGGACCTC CATGGGTGAC
 1001 ATGGTAATGC TGGCCTGCTG GCTGAACCCC AGATTGCCAT GTTCTGTGGC
 TACCATTACG ACCGGACGAC CGACTTGGGG TCTAACGGTA CAAGACACCG
 1051 AGACTGAACA TGACATGAA TGTCAAGAAT GGGAAAGTGGG ATTCAAGATCC
 TCTGACTTGT ACGTGTACTT ACAGGTCTTA CCCTTCACCC TAAGTCTAGG
 1101 ATCAGGGGACCA AAAACCTGCA TTGATACCAA GGAAGGCATC CTGCACTATT
 TAGTCCCTGG TTTTGGACGT AACTATGGTT CCTTCCGTAG GACGTCA
 1151 GCCAAGAAGT CTACCCGAA CTGCAAGATCA CCAATGTGGT AGAACCC
 CGGTTCTCA GATGGGACTT GACGTCTAGT GTTACACCA TCTTCGGTT
 1201 CAACCACTGA CCATCCGAA CTGGTGCAAG CGGGGCGCAG AGCAGTGCA
 GTTGGTCACT GGTAGGTCTT GACCACTTC GCCCCCGCGT TCGTCACGTT
 1251 GACCCATCCC CACTTGTGA TTCCCTACCG CTGCTTAGTT GGTGAGTTA
 CTGGGTAGGG GTGAAACACT AAGGGATGGC GACGAATCAA CCACTCAAAT
 1301 TAAGTGATGC CTTCTCGTT CCTGACAAGT GCAAATTCTT ACACCAAGGAG
 ATTCACTACG GGAAGAGCAA GGACTGTTCA CGTTAAAGAA TGTGGTC
 1351 AGGATGGATG TTGCGAAAC TCATCTTACG TGGCACACCG TCGCCAAAGA
 TCCTACCTAC AAACGTTTG AGTACAAGT ACCGTGTGGC AGCGGTTTCT

FIG.26B

48/76

1401	GACATGCAGT	GAGAAGAGTA	CCAAC TTGCA	TGACTACGGC	ATGTTGCTGC
	CTGTACGTCA	CTCTTCTCAT	GGTTGAACGT	ACTGATGCCG	TACAACGACG
1451	CCTCGGAAAT	TGACAAGTT	CGAGGGGTAG	AGTTTGTTG	TTGCCCACTG
	GGACGCCCTTA	ACTGTTCAAG	GCTCCCCATC	TCAAACACAC	AACGGGTGAC
1501	GCTGAAGAAA	GTGACAATGT	GGATTCTGCT	GATGCGGAGG	AGGATGACTC
	CGACTTCTT	CACTGTTACA	CCTAACAGCA	CTACGCCCTC	TCCTACTGAG
1551	GGATGTCTGG	TGGGGCGGGAG	CAGACACAGA	CTATGCAGAT	GGGAGTGAAG
	CCTACAGACC	ACCCCGCCCTC	GTCTGTGTC	GATACTGCTA	CCCTCACTTC
1601	ACAAAGTAGT	AGAAGTAGCA	GAGGGAGGAAG	AAGTGGCTGA	GGTGGAAAGAA
	TGTTTCATCA	TCTTCATCGT	CTCCTCCTTC	TTCACCGACT	CCACCTTCTT
1651	GAAGAAGCCG	ATGATGACGA	GGACGATGAG	GATGGTGATG	AGGTAGAGGA
	CTCTTCCGGC	TACTACTGCT	CCTGCTACTC	CTACCACTAC	TCCATCTCCT
1701	AGAGGCTGAG	GAACCCCTACG	AAGAACCCAC	AGAGAGAAC	ACCAGCATTG
	TCTCCGACTC	CTTGGGATGC	TTCTTCGGTG	TCTCTCTTG	TGGTCGTAAC
1751	CCACCAACAC	CACCAACACC	ACAGAGTCTG	TGGAAGAGGT	GGTTCGAGTT
	GGTGGTGGTG	GTGGTGGTGG	TGTCAGAC	ACCTTCTCCA	CCAAGCTCAA
1801	CCTACAAACAG	CAGCCAGTAC	CCCTGATGCC	GTTGACAAGT	ATCTCGAGAC
	GGATGTTGTC	GTCGGTCA	GGGACTACGG	CAACTGTTCA	TAGAGCTCTG
1851	ACCTGGGGAT	GAGAATGAAC	ATGCCCATTT	CCAGAAAGCC	AAAGAGAGGC
	TGGACCCCTA	CTCTTACTTG	TACGGTAAA	GGTCTTCGG	TTTCTCTCG
1901	TTGAGGCCAA	GCACCGAGAG	AGAATGTCCC	AGGTCTGAG	AGAATGGGAA
	AACTCCGGTT	CGTGGCTCTC	TCTTACAGGG	TCCAGTACTC	TCTTACCCCTT
1951	GAGGCAGAAC	GTCAAGAAA	GAACCTGCCT	AAAGCTGATA	AGAAGGCAGT
	CTCCGTCTTG	CAGTCGTTT	CTTGAACGGA	TTTCGACTAT	TCTTCCGTCA
2001	TATCCAGCAT	TTCCAGGAGA	AAAGTGAATC	TTTGGAACAG	GAAGCAGCCA
	ATAGGTGCTA	AAGGTCTCT	TTCACCTTAG	AAACCTTGTC	CTTCGTCGGT
2051	ACGAGAGACA	GCAGCTGGTG	GAGACACACA	TGGCCAGAGT	GGAAAGCCATG
	TGCTCTCTGT	CGTCGACCCAC	CTCTGTGTC	ACCGGTCTCA	CCTTCGGTAC
2101	CTCAATGACC	GCCGCCGCCT	GGCCCTGGAG	AACTACATCA	CCGCTCTGCA
	GAGTTACTGG	CGGGCGCGGA	CCGGGACCTC	TTGATGTAGT	GGCGAGACGT
2151	GGCTGTTCT	CCTCGGCCTC	GTCACGTGTT	CAATATGCTA	AAGAAGTATG
	CCGACAAGGA	GGAGCCGGAG	CAGTGCACAA	GTTATACGAT	TTCTTCATAC
2201	TCCGCGCAGA	ACAGAAGGAC	AGACAGCACA	CCCTAAAGCA	TTTCGAGCAT
	AGGCGCGTCT	TGTCTTCCTG	TCTGTCGTG	GGGATTTCGT	AAAGCTCGTA
2251	GTGCGCATGG	TGGATCCCAA	GAAAGCCGCT	CAGATCCGGT	CCCAGGTTAT
	CACCGCGTACC	ACCTAGGGTT	CTTCCGGCGA	GTCTAGGCCA	GGGTCCAATA
2301	GACACACCTC	CGTGTGATTT	ATGAGCGCAT	GAATCAGTCT	CTCTCCCTGC
	CTGTGTGGAG	GCACACTAAA	TACTCGCGTA	CTTAGTCAGA	GAGAGGGACG
2351	TCTACAAACGT	GCCTGCAGTG	GCCGAGGAGA	TTCAGGATGA	AGTTGATGAG
	AGATGTTGCA	CGGACGTCAC	CGGCTCCTCT	AAGTCCTACT	TCAACTACTC
2401	CTGCTTCAGA	AAGAGAAAAA	CTATTCACT	GACGTCTTGG	CCAACATGAT
	GACGAAGTCT	TTCTCGTTT	GATAAGTCTA	CTGCAGAAC	GGTTGTA
2451	TAGTGAACCA	AGGATCAGT	ACGGAAACGA	TGCTCTCATG	CCATCTTGA
	ATCACTTGGT	TCCTAGTCAA	TGCTTTGCT	ACGAGAGTAC	GGTAGAAACT
2501	CCGAAACGAA	AACCACCGTG	GAGCTCCTC	CCGTGAATGG	AGAGTTCAGC
	GGCTTGCTT	TTGGTGGCAC	CTCGAGGAAG	GGCACTTACC	TCTCAAGTCG
2551	CTGGACGATC	TCCAGCGCTG	GCATTCTTT	GGGGCTGACT	CTGTGCCAGC
	GACCTGCTAG	AGGTGGCAC	CGTAAGAAAA	CCCCGACTGA	GACACGGTCG
2601	CAACACAGAA	AACGAAGTTG	AGCCTGTTGA	TGCCCCGCCCT	GCTGCCGACC
	GTTGTGTC	TTGCTTCAAC	TCGGACAACT	ACGGGCGGGGA	CGACGGCTGG
2651	GAGGAAGTGCAC	CACTCGACCA	GGTTCTGGGT	TGACAAATAT	CAAGACGGAG
	CTCCTGACTG	GTGAGCTGGT	CCAAGACCCA	ACTGTTTATA	GTTCTGCC
2701	GAGATCTCTG	AAGTGAATCT	AGATGCAGAA	TTCCGACATG	ACTCAGGATA
	CTCTAGAGAC	TTCACTTAGA	TCTACGTCTT	AAGGCTGTAC	TGAGTCCTAT
2751	TGAAGTTCAT	CATAAAAAT	TGGTGTCTT	TGCAAGAAGAT	GTGGGTTCAA
	ACTTCAAGTA	GTAGTTTTA	ACCACAAGAA	ACGTCTTCTA	CACCCAAGTT
2801	ACAAAGGTGC	AATCATTGGA	CTCATGGTGG	GCGGTGTTGT	CATAGCGACA
	TGTTCCACG	TTAGTAACCT	GAGTACCA	CGCCACAACA	GTATCGCTG

FIG.26C

49/76

2851	G TGATCGTCA	T CACCTGGT	G ATGCTGAAG	A AGAAAGATA	T CATGGAGCC
	CACTAGCAGT	AGTGGAACCA	CTACGACTTC	TTCTTTCTAT	AGTACCTCGG
2901	AGTAGATCCT	AGACTAGAGC	CCTGGAAGCA	TCCAGGAAGT	CAGCCTAAAA
	TCATCTAGGA	TCTGATCTCG	GGACCTTCGT	AGGTCTTCA	GTCGGATTTT
2951	CTGCTTGTAC	CAATTGCTAT	TGTAAAAAGT	GTTGCTTCA	TTGCCAAGTT
	GACGAACATG	GTAAACGATA	ACATTTTCA	CAACGAAAGT	AACGGTTCAA
3001	TGTTTCATGA	CAAAAGCCTT	AGGCATCTCC	TATGGCAGGA	AGAAGCGGAG
	ACAAAGTACT	GTTTTCGAA	TCCGTAGAGG	ATACCGTCCT	TCTTCGCCTC
3051	ACAGCGACGA	AGAGCTCATC	AGAACAGTCA	GACTCATCAA	GCTTCTCTAT
	TGTCGCTGCT	TCTCGAGTAG	TCTTGTCACT	CTGAGTAGTT	CGAAGAGATA
3101	CAAAGCAGTA	AGTAGGCGGC	CGCTCGAGTC	TAGAGGGCCC	GTTTAAACCC
	GTTCGTCAT	TCATCCGGC	GCGAGCTCAG	ATCTCCCCGG	CAAATTGGG
3151	GCTGATCAGC	CTCGACTGTG	CCTTCTAGTT	GCCAGCCATC	TGTTGTTGC
	CGACTAGTCG	GAGCTGACAC	GGAAGATCAA	CGGTCGGTAG	ACAACAAACG
3201	CCCTCCCCG	TGCCTCTT	GACCTTGAA	GTTGCCACTC	CCACTGTCT
	GGGAGGGGGC	ACGGAAGGAA	CTGGGACCTT	CCACGGTAG	GGTGACAGGA
3251	TTCCAATAA	AATGAGGAAA	TTGCATCGCA	TTGTCTGAGT	AGGTGTATT
	AAGGATTATT	TTACTCCTT	AACGTAGCGT	AACAGACTCA	TCCACAGTAA
3301	CTATTCTGGG	GGGTGGGGTG	GGGCAGGACA	GCAAGGGGGA	GGATTGGGAA
	GATAAGACCC	CCCACCCAC	CCCCTCCTGT	CGTTCCCCCT	CCTAACCCCTT
3351	GACAATAGCA	GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	CTTCTGAGGC
	CTGTTATCGT	CCGTACGACC	CCTACGCCAC	CCGAGATACC	GAAGACTCCG
3401	GGAAAAGAAC	AGCTGGGGCT	CTAGGGGTA	TCCCCACGCG	CCCTGTAGCG
	CCTTCTTGGG	TCGACCCCCA	GATCCCCAT	AGGGGTGCGC	GGGACATCGC
3451	GCGCATTAAAG	CGCGGCGGGT	GTGGTGGTTA	CGCGCAGCGT	GACCGCTACA
	CGCGTAATTTC	CGCCGCCCA	CACCACCAAT	CGCGCTCGCA	CTGGCGATGT
3501	CTTGCAGCG	CCCTAGCGCC	CGCTCTTTC	GCTTCTTTC	CTTCCTTCT
	GAACGGTCGC	GGGATCGGG	GCGAGGAAAG	CGAAAGAAGG	GAAGGAAAGA
3551	CGCCACGTT	GCCGGTTTC	CCCGTCAAGC	TCTAAATCGG	GGCATCCCTT
	GCGGTGCAAG	CGGCCGAAAG	GGGCAGTTCG	AGATTTAGCC	CCGTAGGGAA
3601	TAGGGTCCG	ATTTAGTGT	TTACGGCACC	TCGACCCCCA	AAAACTTGAT
	ATCCCAAGGC	TAATCACGA	AATGCCGTGG	AGCTGGGGTT	TTTTGAACTA
3651	TAGGGTGATG	GTTCACGTAG	TGGGCCATCG	CCCTGATAGA	CGGTTTTTCG
	ATCCCACTAC	CAAGTGCATC	ACCCGGTAGC	GGGACTATCT	GCCAAAAAGC
3701	CCCTTGACG	TTGGAGTCCA	CGTTTTAA	TAGTGGACTC	TTGTTCAAA
	GGGAAACTGC	AACCTCAGGT	GCAAGAAATT	ATCACCTGAG	AACAAGGTTT
3751	CTGGAACAAC	ACTCAACCT	ATCTCGGTCT	ATTCTTTGA	TTTATAAGGG
	GACCTTGTG	TGAGTTGGGA	TAGAGCCAGA	TAAGAAAACT	AAATATTCCC
3801	ATTTGGGGA	TTTCGGCTA	TTGGTTAAA	AATGAGCTGA	TTAACAAAA
	AAAACCCCT	AAAGCCGGAT	AACCAATT	TTACTCGACT	AAATTGTTT
3851	ATTTAACGCG	ATTAATTCT	GTTGAATGTG	TGTCACTTAG	GGTGTGAAA
	TAATTGCGC	TTAATTAAGA	CACCTTACAC	ACAGTCATC	CCACACCTT
3901	GTCcccAGGC	TCCCCAGCA	GGCAGAACGTA	TGCAAAGCAT	GCATCTCAAT
	CAGGGGTCCG	AGGGGTCCGT	CCGTCCTCAT	ACGTTTCGTA	CGTAGAGTTA
3951	TAGTCAGCAA	CCAGGTGTGG	AAAGTCCCCA	GGCTCCCCAG	CAGGCAGAAG
	ATCAGTCGTT	GGTCCACACC	TTTCAGGGT	CCGAGGGGTC	GTCCGCTTTC
4001	TATGCAAAGC	ATGCATCTCA	ATTAGTCAGC	AACCATAGTC	CCGCCCCCTAA
	ATACGTTTCG	TACGTAGAGT	TAATCAGTCG	TTGGTATTCAG	GGCGGGGGATT
4051	CTCGCCCCAT	CCCGCCCCCA	ACTCCGCCCCA	GTTCGGCCCCA	TTCTCCGCC
	GAGGCAGGGTA	GGGCGGGGAT	TGAGGCGGGT	CAAGGCAGGGT	AAGAGGGGGG
4101	CATGGCTGAC	TAATTTTTT	TATTTATGCA	GAGGCCAGGG	CCGCCTCTGC
	GTACCGACTG	ATAAAAAAA	ATAAATACGT	CTCCGGCTCC	GGCGGAGACG
4151	CTCTGAGCTA	TTCCAGAAGT	AGTGAGGAGG	CTTTTTGGA	GGCCTAGGCT
	GAGACTCGAT	AAGGTCTTCA	TCACTCCTCC	GAAAAAAACCT	CCGGATCCGA
4201	TTTGCAAAAA	GCTCCCGGGA	GCTTGATAT	CCATTTTCGG	ATCTGATCAG
	AAACGTTTT	CGAGGGCCCT	CGAACATATA	GGTAAAAGCC	TAGACTAGTC

FIG.26D

50/76

4251	CACGTGTTGA	CAATTAATCA	TCGGCATAGT	ATATCGGCAT	AGTATAATAC
	GTGCACAAC	GTAAATTAGT	AGCCGTATCA	TATAGCCGTA	TCATATTATG
4301	GACAAGGTGA	GGAACTAAC	CATGCCAAG	TTGACCAGTG	CCGTTCCGGT
	CTGTTCACT	CCTTGATTG	GTACCGGTT	AACTGGTCAC	GGCAAGGCCA
4351	GCTCACCGCG	CGCGACGTCG	CCGGAGCGGT	CGAGTTCTGG	ACCGACCGGC
	CGAGTGGCGC	GCGCTGCAGC	GGCCTCGCCA	GCTCAAGACC	TGGCTGCCG
4401	TCGGGTTCTC	CCGGGACTTC	GTGGAGGACG	ACTTCGCCG	TGTGGTCCGG
	AGCCCAAGAG	GGCCCTGAAG	CACCTCTGC	TGAAGCGGCC	ACACCAAGGCC
4451	GACGACGTGA	CCCTGTTCAT	CAGCGCGTC	CAGGACGAGG	TGGTGCAGGA
	CTGCTGCACT	GGGACAAGTA	GTCGCGCCAG	GTCTCTGGTC	ACCACGGCCT
4501	CAACACCCCTG	GCCTGGGTG	GGGTGCGCGG	CCTGGACGAG	CTGTACGCCG
	GTTGTGGGAC	CGGACCCACA	CCCACGCGCC	GGACCTGCTC	GACATGCCG
4551	AGTGGTCGGA	GGTGTGTC	ACGAACTTCC	GGGACGCCTC	CGGGCCGGCC
	TCACCAGCCT	CCAGCACAGG	TGCTTGAAGG	CCCTGCGGAG	GCCCCGGCCGG
4601	ATGACCGAGA	TGCGCGAGCA	GCCGTTGGGG	CGGGAGTTCG	CCCTGCGCGA
	TACTGGCTCT	AGCGCTCTG	CGGCACCCCC	GCCCTCAAGC	GGGACCGCGCT
4651	CCCGGGCGGC	AACTGCGTC	ACTTCGTTGC	CGAGGAGCAG	GAUTGACACG
	GGGCGGGCGC	TTGACGCAAG	TGAAGCACCG	GCTCCTGTC	CTGACTGTGC
4701	TGCTACGAGA	TTTCGATTCC	ACCGCCGCCT	TCTATGAAAG	GTTGGGCTTC
	ACGATGCTCT	AAAGCTAAGG	TGGCGCGGA	AGATACTTTC	CAACCCGAAG
4751	GGATCGTT	TCGGGACGC	CGGCTGGATG	ATCCTCCAGC	CGGGGGATCT
	CCTTAGCAAA	AGGCCCTGCG	GCCGACCTAC	TAGGAGGTG	CGCCCCTAGA
4801	CATGCTGGAG	TTCTCGCC	ACCCCAACTT	GTTTATTGCA	GCTTATAATG
	GTACGACCTC	AAGAAGCGGG	TGGGGTTGAA	CAAATAACGT	CGAATATTAC
4851	GTTACAAATA	AAGCAATAGC	ATCACAAATT	TCACAAATAA	AGCATTTTTT
	CAATGTTTAT	TTCGTTATCG	TAGTGTAA	AGTGTAA	TCGTAaaaaaa
4901	TCACTGCATT	CTAGTTGTGG	TTTGTCCAAA	CTCATCAATG	TATCTTATCA
	AGTGACGTA	GATCAACACC	AAACAGGTTT	GAGTAGTTAC	ATAGAATAGT
4951	TGTCTGTATA	CGCTCGACCT	CTAGCTAGAG	CTTGGCGTAA	TCATGGTCAT
	ACAGACATAT	GGCAGCTGG	GATCGATCTC	GAACCGCATT	AGTACCAAGTA
5001	AGCTGTTCC	TGTGTAAAT	TGTTATCCGC	TCACAATTCC	ACACAAACATA
	TCGACAAAGG	ACACACTTA	ACAATAGGCG	AGTGTAAAGG	TGTGTTGTAT
5051	CGAGCCGGAA	GCTAAAGTG	TAAAGCCTGG	GGTGCCTAAT	GAGTGAGCTA
	GCTGGCCTT	CGTATTTAC	ATTTCGGACC	CCACGGATT	CTCACTCGAT
5101	ACTCACATTA	ATTGCGTTGC	GCTCACTGCC	CGCTTTCCAG	TCGGGAAACC
	TGAGTGTAAAT	TAAGCGAACG	CGAGTGAAGG	GCGAAAGGTC	AGCCCTTTGG
5151	TGTCGTGCCA	GCTGCATTAA	TGAATCGGCC	AACCGCGGG	GAGAGCGGT
	ACAGCACGGT	CGACGTAATT	ACTTAGCCG	TTGCGCGCCC	CTCTCCGCCA
5201	TTGCGTATTG	GGCGCTCTTC	CGCTTCTCG	CTCACTGACT	CGCTGCGCTC
	AACGCATAAC	CCGCGAGAAG	GCGAAGGAGC	GAGTGAATG	GCGACGCGAG
5251	GGTCGTTCGG	CTGCGCGAG	CGGTATCAGC	TCACTCAAAG	CGGGTAATAC
	CCAGCAAGCC	GACGCCGCTC	GCCATAGTCG	AGTGAAGTTTC	CGCCATTATG
5301	GGTTATCCAC	AGAATCAGGG	GATAACGAG	GAAAAGAACAT	GTGAGCAAAA
	CCAATAGGTG	TCTTAGTCCC	CTATTGCGTC	CTTCTTGTG	CACTCGTTT
5351	GGCAGCAAA	AGGCCAGGAA	CGTAAAAAG	GCCCGCGTGC	TGGCGTTTTT
	CCGGTCGTT	TCCGGTCTT	GGCATTTTC	CGCGCGAACG	ACCGCAaaaaa
5401	CCATAGGCTC	CGCCCCCTG	ACGAGCATCA	AAAAATCGA	CGCTCAAGTC
	GGTATCCGAG	GCGGGGGGAC	TGCTCGTAGT	GTTTTAGCT	GCGAGTTCA
5451	AGAGGTGGCG	AAACCCGACA	GGACTATAAA	GATACCAGGC	GTTTCCCCCT
	TCTCCACCGC	TTTGGGCTGT	CCTGATATT	CTATGGTCCG	CAAAGGGGGA
5501	GGAAGCTCCC	TCGTGCGCTC	TCCTGTTCCG	ACCCTGCCG	TTACCGGATA
	CCTTCGAGGG	AGCACGCGAG	AGGACAAGGC	TGGGACGGCG	AATGGCTAT
5551	CCTGTCCGCC	TTTCTCCCTT	CGGGAAGCGT	GGCGCTTTCT	CAATGCTCAC
	GGACAGGGCGG	AAAGAGGGAA	GCCCTCGCA	CCGCAGAAAGA	GTTACGAGTG
5601	GCTGTAGGTA	TCTCGATTCG	GTGTAGGTCG	TTGCTCCAA	GCTGGGCTGT
	CGACATCCAT	AGAGTCAGC	CACATCCAGC	AAGCGAGGTT	CGACCCGACA

FIG.26E

51/76

5651	GTGCACGAAC	CCCCCGTTCA	GCCCGACCGC	TGCGCCTTAT	CCGGTAACTA
	CACGTGCTTG	GGGGGCAAGT	CGGGCTGGCG	ACGCGGAATA	GGCCATTGAT
5701	TCGTCTTGAG	TCCAACCCGG	TAAGACACGA	CTTATCGCCA	CTGGCAGCAG
	AGCAGAACTC	AGGTTGGCC	ATTCTGTGCT	GAATAGCGGT	GACCGTCGTC
5751	CCACTGGTAA	CAGGATTAGC	AGAGCAGGGT	ATGTAGGCGG	TGCTACAGAG
	GGTGACCATT	GTCCTAAATCG	TCTCGCTCCA	TACATCCGCC	ACGATGTCTC
5801	TTCTTGAAGT	GGTGGCTAA	CTACCGCTAC	ACTAGAAGGA	CAGTATTGG
	AAGAACCTCA	CCACCGGATT	GATGCCGATG	TGATCTTCCT	GTCATAAACC
5851	TATCTGCGCT	CTGCTGAAGC	CAGTTACCTT	CGGAAAAAGA	GTTGGTAGCT
	ATAGACGCGA	GACGACTTCG	GTCAATGGAA	GCCTTTTCT	CAACCATCGA
5901	CTTGATCCGG	CAAACAAACC	ACCGCTGGTA	GCGGTGGTTT	TTTTGTTTGC
	GAACTAGGCC	GTGTTGTTGG	TGGCGACCAT	CGCCACCCAA	AAAACAAACG
5951	AAGCAGCAGA	TTACGCGCAG	AAAAAAAGGA	TCTCAAGAAG	ATCCTTTGAT
	TTCGTCGCT	TTTTTTTCT	AGAGTTCTTC	TAGGAAACTA	
6001	CTTTCTACG	GGGTCTGACG	CTCAGTGGAA	CGAAAAACTCA	CGTTAAGGGA
	GAAAAGATGC	CCCAGACTGC	GAGTCACCTT	GCTTTTGAGT	GCAATTCCCT
6051	TTTGGTCAT	GAGATTATCA	AAAAGGATCT	TCACCTAGAT	CCTTTAAAT
	AAAACCGATA	CTCTAATAGT	TTTTCCTAGA	AGTGGATCTA	GGAAAATTAA
6101	TAAAATGAA	GTTTTAAATC	AATCTAAAGT	ATATATGAGT	AAACTTGGTC
	ATTTTACTT	CAAAATTAG	TTAGATTCA	TATATACTCA	TTTGAACCAG
6151	TGACAGTTAC	CAATGTTAA	TCAGTGAGGC	ACCTATCTCA	GCGATCTGTC
	ACTGTCAATG	GTTACGAATT	AGTCACTCCG	TGGTAGAGT	CGCTAGACAG
6201	TATTCGTT	ATCCATAGTT	GCCTGACTCC	CGCTCGTGT	GATAACTACG
	ATAAAGCAAG	TAGGTATCAA	CGGACTGAGG	GGCAGCACAT	CTATTGATGC
6251	ATACGGGAGG	GCTTACCATC	TGGCCCCAGT	GCTGCAATGA	TACCGCGAGA
	TATGCCCTCC	CGAATGGTAG	ACCGGGGTCA	CGACGTTACT	ATGGCGCTCT
6301	CCCACGCTCA	CCGGCTCCAG	ATTATCAGC	AATAAAACCAG	CCAGCCGGAA
	GGGTGCGAGT	GGCCGAGGTC	TAATAGTCG	TTATTGGTC	GGTCGGCCTT
6351	GGGCCGAGCG	CAGAAGTGGT	CCTGCAACTT	TATCCGCTC	CATCCAGTCT
	CCCGGCTCGC	GTCTTCACCA	GGACGTTGAA	ATAGGCGGAG	GTAGGTCAAG
6401	ATTAATTGTT	GCCGGGAAGC	TAGAGTAAGT	AGTTCGCCAG	TTAATAGTTT
	TAATTAACAA	CGGCCCTTCG	ATCTCATTCA	TCAAGCGGTC	AATTATCAA
6451	GCGCAACGTT	GTTGCCATTG	CTACAGGCAT	CGTGGTGTCA	CGCTCGTCGT
	CGCGTTGCAA	CAACGGTAAC	GATGTCGTA	GCACCACAGT	GCGAGCAGCA
6501	TTGGTATGGC	TTCATTCAAGC	TCCGGTTCCC	AACGATCAAG	GCGAGTTACA
	AACCATAACCG	AAGTAAGTCG	AGGCCAAGGG	TTGCTAGTT	CGCTCAATGT
6551	TGATCCCCCA	TGTTGTGCAA	AAAAGCGGTT	AGCTCCTTCG	GTCCTCCGAT
	ACTAGGGGGT	ACAACACGTT	TTTTCGCCAA	TCGAGGAAGC	CAGGAGGCTA
6601	CGTTGTCAGA	AGTAAGTTGG	CCGCAGTGT	ATCACTCATG	GTTATGGCAG
	GCAACAGTCT	TCATTCAACC	GGCCTCACAA	TAGTGGAGTAC	CAATACCGTC
6651	CACTGCATAA	TTCTCTTACT	GTCATGCCAT	CCGTAAGATG	CTTTTCTGTG
	GTGACGTATT	AAGAGAATGA	CAGTACGGTA	GGCATTCTAC	GAAAAGACAC
6701	ACTGGTGAGT	ACTCAACCAA	GTCATTCTGA	GAATAGTGT	TGCGGCAGCC
	TGACCACTCA	TGAGTTGGTT	CAGTAAGACT	CTTATCACAT	ACGCCGCTGG
6751	GAGTTGCTCT	TGCCCCGGCGT	CAATACGGGA	TAATACCGCG	CCACATAGCA
	CTCAACGAGA	ACGGGGCCGCA	GTTATGCCCT	ATTATGGCGC	GGTGTATCGT
6801	GAACTTTAAA	AGTGCTCATC	ATTGGAAAAC	GTTCCTCGGG	GCGAAAACTC
	CTTGAAATT	TCACGAGTAG	TAACCTTTG	CAAGAAGCCC	CGCTTTGAG
6851	TCAAGGATCT	TACCGCTGTT	GAGATCCAGT	TCGATGTAAC	CCACTCGTC
	AGTTCCCTAGA	ATGGCGACAA	CTCTAGGTCA	AGCTACATTG	GGTGAGCAG
6901	ACCCAACGTGA	TCTTCAGCAT	CTTTTACTTT	CACCAGCGTT	TCTGGGTGAG
	TGGGTTGACT	AGAAGTGGTA	GAAAATGAAA	GTGGTCGCAA	AGACCCACTC
6951	CAAAAACAGG	AAGGCAAAT	GCCGCAAAAA	AGGGAATAAG	GGCGACACGG
	GTTTTTGTCC	TTCCGTTTTA	CGGCGTTTTT	TCCCTTATT	CCGCTGTGCC
7001	AAATGTTGAA	TACTCATACT	CTTCCCTTTT	CAATATTATT	GAAGCAATTAA
	TTTACAACCTT	ATGAGTATGA	GAAGGAAAAA	GTATAATAA	CTTCGTAAT

FIG.26F

52/76

7051 TCAGGGTTAT TGTCTCATGA GCGGATAACAT ATTTGAATGT ATTTAGAAAA
AGTCCAATA ACAGAGTAAC CGCCTATGTA TAAACTTACA TAAATCTTTT
7101 ATAAACAAAT AGGGGTTCCG CGCACATTTC CCCGAAAAGT GCCACCTGAC
TATTGTATA TCCCCAAGGC CGGTGTAAAG GGGCTTTCA CGGTGGACTG
7151 GTC
CAG

FIG.26G

53/76

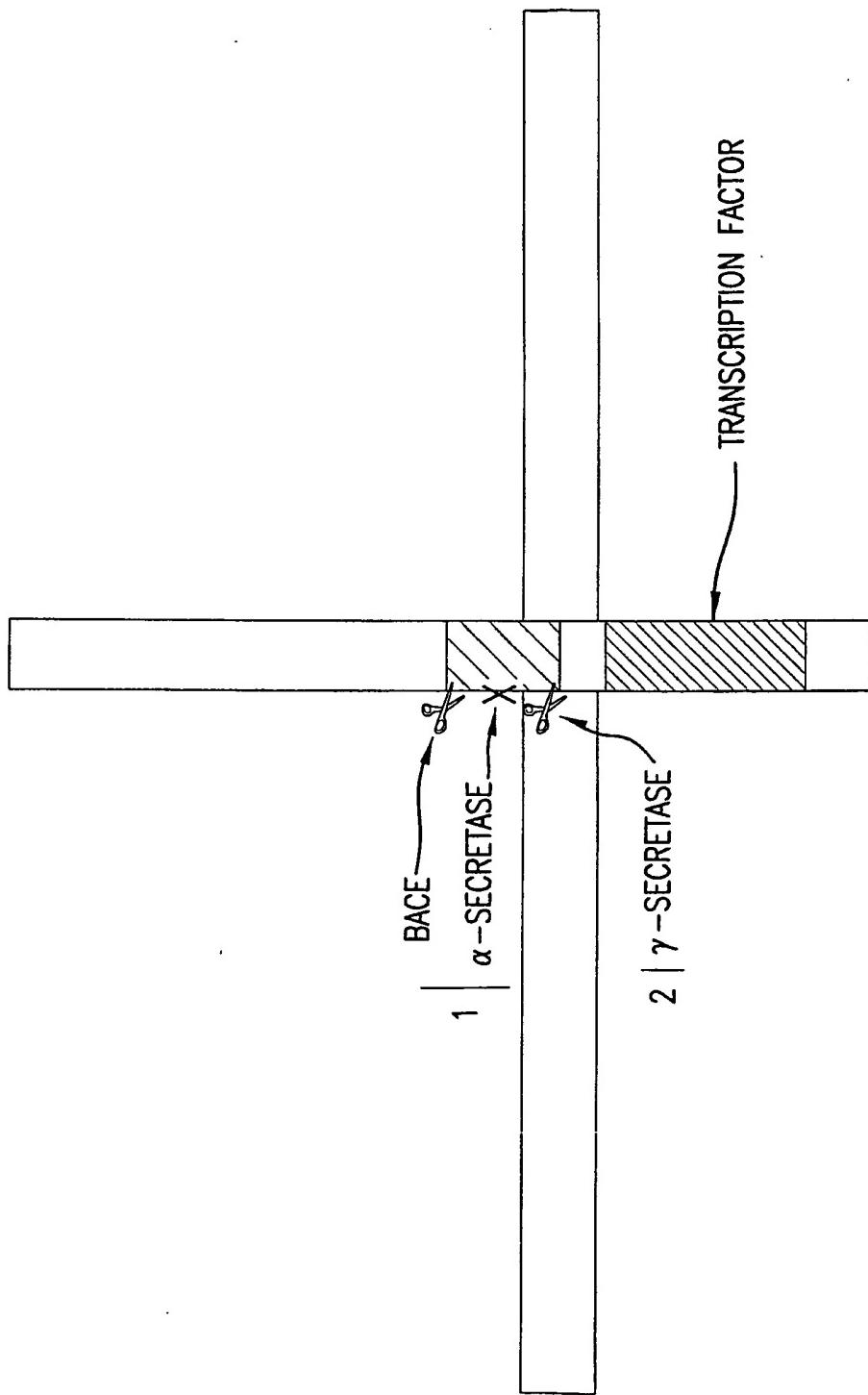


FIG.27A

54/76

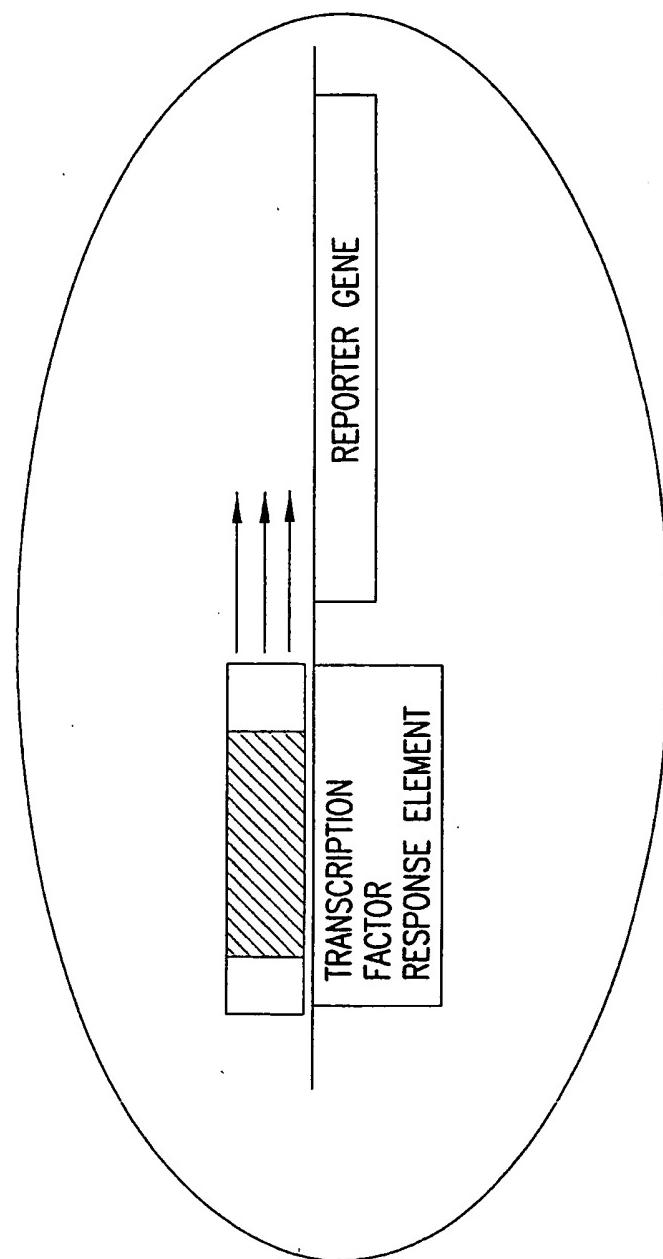


FIG. 27B

55/76

DNA sequence of APP(1-651)NFEV, K612V-TATexon1(M1L) APP (664-695)
(SEQ ID NO: 23)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTG CCTACCGCTG
351 CTTAGTTGGT GAGTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GCGGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTACA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACACC ACCATTGCCA CCACCAACAC CACCACACAC GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCCTCGT CGGCCTCGTC ACGTGTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGACAGA CAGCACACCC
1301 TAAAGCATT CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

FIG.28A

56/76

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCA GTGGCC GAGGAGATT
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAACCTTGAG AGTGGAAATT
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAAATT
2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC
2051 CAGGAAGTCA GCCTAAAAC GCTTGTACCA ATTGCTATTG TAAAAAGTGT
2101 TGCTTTCATT GCCAAGTTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA
2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA
2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC
2251 GCAGACGCCG CTGTCACCCC AGAGGGAGCGC CACCTGTCCA AGATGCAGCA
2301 GAACGGCTAC GAAAATCCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT
2351 AG

FIG.28B

57/76

(SEQ ID NO: 24)

Amino acid sequence of APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695)

m1pg1a1111aawtaralevptdgnag11aepqiamfcgr1nmhmnnvqngkwdsdpsgtktcidtkegilqycq
evyapelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckf1hgermdvceth1h
whtvaketcsekstn1hdym11pcgidkfrgvefvccplaeesdnvdssadaeeddsdvwwggadtdyadgs
1 edkvvevaeeeevaeveeeeadddeddedgdeveeeaeepyeeaterttsiattttttesveevrvpttaastpd
avdkyletpgdenehahfqkakerleakhrermsqvmoreeweaerqaknlpkadkkaviqhfkqekvesleqe
aanerqq1vethmarveam1ndrrrlalenyitalqavpprprhvfnm1kkyvraeqkdrqht1khfehvrnvd
pkkaaqirsqvmtlrviyermnqs1s1lynvpavaeeiqdevdellqkeqnysddvlanmisseprisygndal
mps1tetkttvel1pvngefs1dd1qpwhsfadsvpantenevevpdarpaadrg1ttrpgsg1tnikteeisev
2 3 4 5
nfevefrhdsgyevhhqylvffaedvgsnkaiig1mvgyyiatvivit1vm1kkklgtelgstspwns
6
ad1lepvdpr1epwkhpqspktactncyckccfhcqvcf1kalqisygrkrrrrahqnsqthqas1skq
7 8
rissstvaaadaavtpeerh1skmqnqyenptykffeqmgn

FIG.29

58/76

DNA sequence of APP(1-651)NFEV, K612V-GAL4VP16(delMet) APP (664-695)

(SEQ ID NO: 25)

1 ATGCTGCCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACCTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCGAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTG CCTACCGCTG
351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCACTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGCATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCAGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACCAACC AGCATTGCCA CCACCACCA CACCACACA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAAATCTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCCT CCTGGAGAAC

FIG.30A

59/76

1251 TATGCTAAAG AAGTATGTCC GCGCAGAAC AAGGGACAGA CAGCACACCC
1301 TAAAGCATT CGAGCATGTG CGCATGGTGG ATCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCA GTGGCC GAGGAGATT
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAAC TCGACCAGGT TCTGGGT.TGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAACTTGA AGTGGAAATT
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAGTATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGCG
1901 GTGTTGTCAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATT
2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTTGCC
2051 GACTAAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT
2101 CTGAAGAAC AACTGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC
2151 GCTGACTAGG GCACATCTGA CAGAAGTGGATCAAGGCTA GAAAGACTGG
2201 AACAGCTATT TCTACTGATT TTTCTCGAG AAGACCTTGA CATGATTTG
2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT
2301 ACAAGATAAT GTGAATAAAG ATGCCGTAC AGATAGATTG GCTTCAGTGG
2351 AGACTGATAT GCCTCTAAC A TTGAGACAGC ATAGAATAAG TGCGACATCA
2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG
2451 AATTCCCCGGG GATCTGGCCC CCCCACCGA TGTCAGCCTG GGGGACGAGC
2501 TCCACTTGA CGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC
2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.30B

60/76

2601 TACCCCCCAC GACTCCGCC CCTACGGCGC TCTGGATATG GCCGACTTCG
2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG
2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG
2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT
2801 TCTTGAGCA GATGCAGAAC TAG

FIG.30C

61/76

(SEQ ID NO: 26)

Amino acid sequence of APP(1-651)NFEV, K612V, GAL4-VP16(delM1) APP (664-695)

m1pglal111aawtaralevptdgnag11aepqiamfcgr1nmhmnnvqngkwdsdpsgtktcidtkegilqycq
evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckf1hqermvdvceth1h
whtvaketcsekstn1hdygml1pcgidkfrgvefvccplaeesdnvdssadaeeddsdvwggadtdyadgs
1
edkvvevaeeeevaeveeeeadddeddedgdeveeeaeypyeeaterttsiatttttesveevrvpttaastpd
avdkyletpgdenehahfqkakerleakhrermsqvmrewaeaerqakn1pkadkkaviqhffqekvesleqe
aanerqqlvethmarveam1ndrrrlaleniyitalqavpprprhvfnm1kkyvraeqkdrqht1khfehvrmd
pkkaaqirsqvmtth1rviyermnqs1s1lynvpavaeeiqdevde11qkeqnyssddvlannmisepri sygndal
mps1tetkttvel1pvngefs1dd1qpwhsfadsvpantenevepvdarpaadrg1ttrpgsg1tnikteeisev
2 3 4 5
nfevefrhdsgyevhhqy1vffaedvgsnkgaig1mvgyyiatyivit1vmlkkkk1telgstspwwwns
ad1k11ssieqacd1c1kk1kcskekpkcakc1knrwecrysptkrsp1trah1tevesrlerleq1f11ifpred1d
6
milkm1ds1qdikalltg1fvqdnvnkdavtdrlasvetdmpl1rqhrisatssseessnkqqrq1tvsgipgdlapp
tdvs1qde1h1dqedvamahad1ddf1dml1qdgdpqgftphdsapya1dmadf1fegmftdal1gidey
7 8
ggd1qhs1gaaadaaytpeerh1skmqngnyenptykffeqmgn

FIG.31

62/76

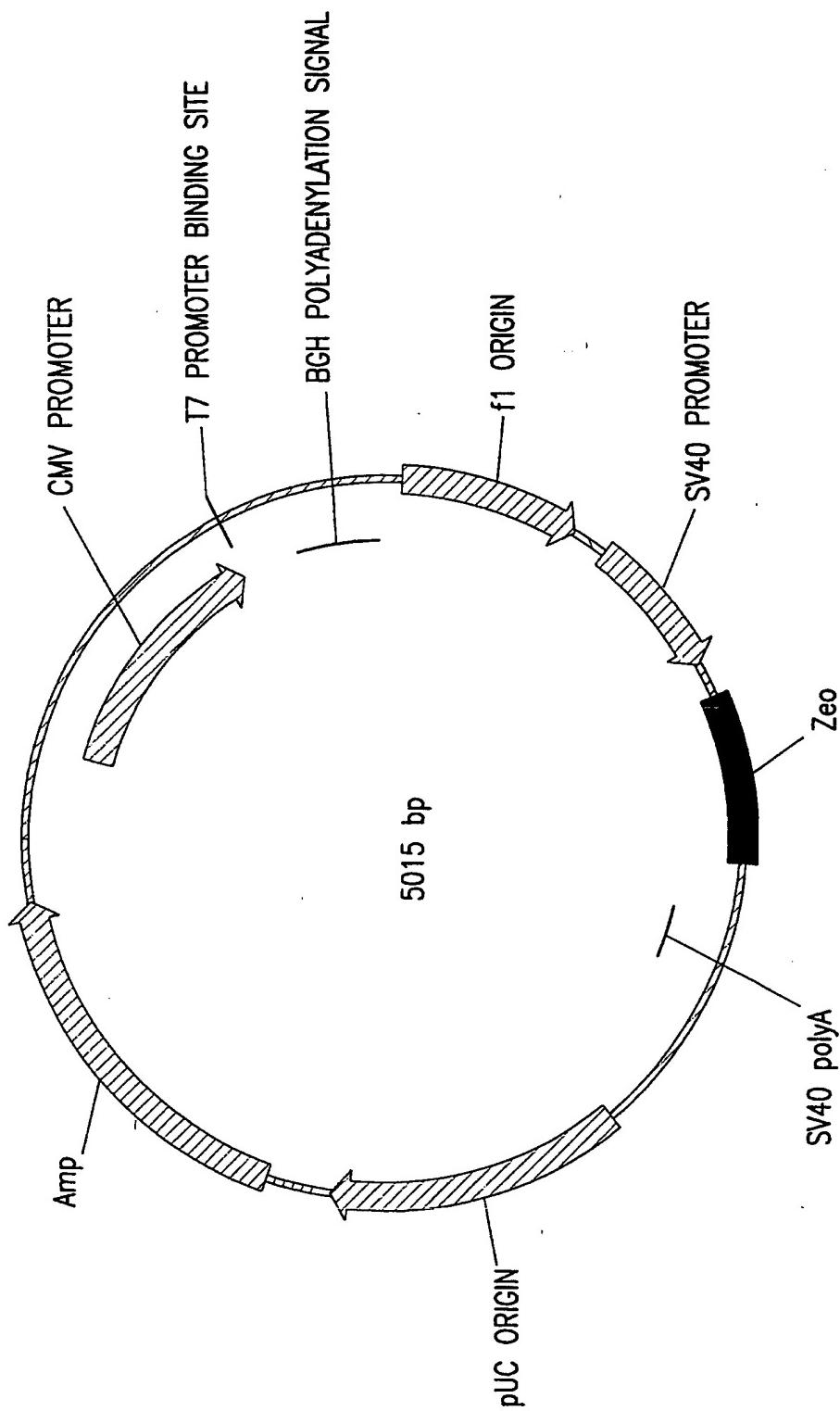


FIG. 32A

63/76

(SEQ ID NO: 27 AND 28)

1 GACGGATCGG GAGATCTCCC GATCCCSTAT GGTCGACTCT CAGTACAATC
 CTGCCTAGCC CTCTAGAGGG CTAGGGGATA CCAGCTGAGA GTCATGTAG
 51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT
 ACGAGACTAC GGCCTATCAA TTCGGTCATA GACGAGGGAC GAACACACAA
 101 GGAGGTCGCT GAGTAGTGC CGAGCAAAAT TTAAGCTACA ACAAGGCAAG
 CCTCCAGCGA CTCATCACGC GCTCGTTTA ATTTCGATGT TGTTCCGTT
 151 GCTTGACCGA CAATTGCAATG AAGAATCTGC TTAGGGTTAG GCGTTTGCG
 CGAACCTGGCT GTTAACGTAC TTCTTAGACG AATCCCAATC CGCAAAACGC
 201 CTGCTTCGCG ATGTACGGC CAGATATAACG CGTTGACATT GATTATTGAC
 GACGAAGCGC TACATGCCCG GTCTATATGC GCAACTGTAA CTAATAACTG
 251 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATA
 ATCAATAATT ATCATTAGTT AATGCCCAAG TAATCAAGTA TCAGGGTATAT
 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG
 ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC
 351 CCCAACGACC CCCGCCATT GACGTCATAA ATGACGTATG TTCCCATAGT
 GGGTTGCTGG GGGCGGGTAA CTGCACTTAT TACTGCATAC AAGGGTATCA
 401 AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAC TATTTACGGT
 TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTG ATAAATGCCA
 451 AAACGTGCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC
 TTTGACGGGT GAACCGTCAT GTAGTTACA TAGTATACGG TTCATGCGGG
 501 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA
 GGATAACTGC AGTTACTGCC ATTACCGGG CGGACCGTAA TACGGGTCA
 551 CATGACCTTA TGGGACTTTCTACTTGGCA GTACATCTAC GTATTAGTCA
 GTACTGGAAT ACCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT
 601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGG
 AGCGATAATG GTACCAACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 651 TAGCGGTTTG ACTCACGGGG ATTTCGAAGT CTCCACCCCA TTGACGTCAA
 ATCGCCAAAC TGAGTGCCTT TAAAGGTTCA GAGGTGGGGT AACTGCAGTT
 701 TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTCCA AAATGTCGTA
 ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTACAGCAT
 751 ACAACTCCGC CCCATTGACG CAAATGGCG GTAGGCGTGT ACGGTGGAG
 TGTGAGGCG GGGTAACTGC GTTACCCGC CATCCGCACA TGCCACCTC
 801 GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA CTGCTTACTG
 CAGATATATT CGTCTCGAGA GACCGATTGA TCTCTTGGGT GACGAATGAC
 851 GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
 CGAATAGCTT TAATTATGCT GAGTGTATC CCTCTGGGTT CGACCGATCG
 901 GTTTAAACTT AAGCTTGGTA CCGAGCTCGG ATCCACTAGT CCAGTGTGGT
 CAAATTGAA TTGGAACCAT GGCTCGAGCC TAGGTGATCA GGTACACCCA
 951 GGAATTCTGC AGATATCCAG CACAGTGGCG GCGCCTCGAG TCTAGAGGGC
 CCTTAAGACG TCTATAGTC GTGTCACCGC CGGCGAGCTC AGATCTCCG
 1001 CCGTTAAAC CCGCTGATCA GCCTCGACTG TGCCCTCTAG TTGCCAGCCA
 GGCAATTGGT GGCGACTAGT CGGAGCTGAC ACGGAAGATC AACGGTCGGT
 1051 TCTGTTGTTT GCCCCTCCCC CGTGCCTTCC TTGACCCCTGG AAGGTGCCAC
 AGACAAACAAA CGGGGAGGGG GCACCGAAGG AACTGGGACC TTCCACGGTG
 1101 TCCCACGTGC CTTTCTAAT AAAATGAGGA ATTGCATCG CATTGTCGTA
 AGGGTGACAG GAAAGGATTA TTTTACTCCT TTAACGTAGC GTAACAGACT
 1151 GTAGGGTGTCA TTCTATTCTG GGGGGTGGGG TGGGGCAGGA CAGCAAGGGG
 CATCCACAGT AAGATAAGAC CCCACCCACCC ACCCCGTCCT GTCGTTCCCC

FIG.32B

64/76

1201 GAGGATTGGG AAGACAATAG CAGGCATGCT GGGGATGCGG TGGGCTCTAT
 CTCCTAACCC TTCTGTTATC GTCCGTACGA CCCCTACGCC ACCCGAGATA
 1251 GGCTTCTGAG GCGGAAAGAA CCAGCTGGGG CTCTAGGGGG TATCCCCACG
 CGGAAGACTC CGCCTTTCTT GGTGACCCCC GAGATCCCCC ATAGGGGTGC
 1301 CGCCCTGTAG CGGCGCATTA AGCGCGGCCGG GTGTGGTGGT TACGCGCAGC
 GCGGGACATC GCGCGTAAT TCGCGCCGCC CACACCACCA ATGCGCGTCG
 1351 GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCGCCTCCTT TCGCTTTCTT
 CACTGGCGAT GTGAACGGTC CGGGGATCGC GGGCGAGGAA AGCGAAAGAA
 1401 CCCTTCCTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC
 GGGAGGAAA GAGCGGTGCA AGCGGCCGAA AGGGGGCAGTT CGAGATTTAG
 1451 GGGGCATCCC TTTAGGGTC CGATTTAGTG CTTTACGGCA CCTCGACCCC
 CCCCGTAGGG AAATCCCAAG GCTAAATCAC GAAATGCCGT GGAGCTGGGG
 1501 AAAAAACTTG ATTAGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA
 TTTTTGAAC TAATCCCAC ACCAAGTGCA TCACCCGGTA GCGGGACTAT
 1551 GACGGTTTT CGCCCTTGA CGTTGGAGTC CACGTTCTT AATAGTGGAC
 CTGCCAAAAA GCGGGAAACT GCAACCTCAG GTGCAAGAAA TTATCACCTG
 1601 TCTTGTTCGA AACTGGAAACA ACACCTCAACC CTATCTCGGT CTATTCTTT
 AGAACAAAGGT TTGACCTTGT TGTGAGTTGG GATAGAGCCA GATAAGAAAA
 1651 GATTTATAAG GGATTTGGG GATTTGGGCC TATTGGTTAA AAAATGAGCT
 CTAATATTCTT CCTAAACCC CTAAGCCGG ATAACCAATT TTTTACTCGA
 1701 GATTTAACAA AAATTTAACG CGAATTAATT CTGTGGAATG TGTGTCAGTT
 CTAATTGTT TTTAAATTGC GCTTAATTAA GACACCTTAC ACACAGTCAA
 1751 AGGGTGTGGA AAGTCCCCAG GCTCCCCAGG CAGGCAGAAG TATGCAAAGC
 TCCCACACCT TTCAGGGTGC CGAGGGTCC GTCCGTCTTC ATACGTTCG
 1801 ATGCATCTCA ATTAGTCAGC AACCAAGGTGT GGAAAGTCCC CAGGCTCCCC
 TACGTAGAGT TAATCAGTCG TTGGTCCACA CTTTCAGGG GTCCGAGGGGG
 1851 AGCAGGCAGA AGTATGCAAA GCATGCATCT CAATTAGTCA GCAACCATAG
 TCGTCCGTCT TCATACGTTT CGTACGTAGA GTTAATCAGT CGTTGGTATC
 1901 TCCCACACCT AACTCCGCC ATCCCGCCCC TAACTCCGCC CAGTTCCGCC
 AGGGCGGGGA TTGAGGGGG TAGGGCGGGG ATTGAGGCAGG GTCAAGGCAGG
 1951 CATTCTCCGC CCCATGGCTG ACTAATTCTT TTTATTATG CAGAGGCCGA
 GTAAGAGGCG GGGTACCGAC TGATTAAGAA AAATAAATAC GTCTCCGGCT
 2001 GGCGCCCTCT GCCTCTGAGC TATTCAGAA GTAGTGAGGA GGCTTTTTG
 CCGCGGGAGA CGGAGACTCG ATAAGGTCTT CATCACTCCT CCGAAAAAAC
 2051 GAGGCCTAGG CTTTGCAAA AAGCTCCCAG GAGCTGTAT ATCCATTTC
 CTCGGATCC GAAAACGTTT TTCGAGGGCC CTCGAACATA TAGGTAAG
 2101 GGATCTGATC AGCACGTGTT GACAATTAAT CATCGGCATA GTATATCGC
 CCTAGACTAG TCGTGCACAA CTGTTAATTA GTAGCCGTAT CATATAGCCG
 2151 ATAGTATAAT ACGACAAGGT GAGGAACCTAA ACCATGGCCA AGTTGACCAG
 TATCATATTA TGCTGTTCCA CTCCTTGATT TGGTACCGGT TCAACTGGTC
 2201 TGCGCTTCCG GTGCTCACCG CGCGCGACGT CGCCGGAGCG GTCGAGTTCT
 ACGGCAAGGC CACGAGTGGC GCGCGCTGCA GCGGCCTCGC CAGCTCAAGA
 2251 GGACCGACCG GCTCGGGTTC TCCCGGGACT TCGTGGAGGA CGACTTCGCC
 CCTGGCTGGC CGAGCCCAAG AGGGCCCTGA AGCACCTCCT GCTGAAGCGG
 2301 GGTGTGGTCC GGGACGACGT GACCCCTGTT ATCAGCGCGG TCCAGGACCA
 CCACACCAAGG CCCTGCTGCA CTGGGACAAG TAGTCGCGCC AGGTCCCTGGT
 2351 GGTGGTGCCG GACAACACCC TGGCCTGGGT GTGGGTGCGC GGCCTGGACG
 CCACCAACGGC CTGTTGTGGG ACCGGACCCA CACCCACGCG CGGGACCTGC

FIG.32C

65/76

2401 AGCTGTACGC CGAGTGGTCG GAGGTCGTGT CCACGAAC TT CCGGGACGCC
 TCGACATGCG GCTCACCA CACA GGTGTTGAA GGCCCTGCGG
 2451 TCCGGGCCGG CCATGACCGA GATCGGCAG CAGCCGTGGG GGCAGGAGTT
 AGGCCCGGCC GGTACTGGCT CTAGCCGCTC GTCGGCACCC CCGCCCTCAA
 2501 CGCCCTGC GC GACCCGGCG GCAACTGC GT GCACCTCGT GCGAGGAGC
 GCGGGACGCG CTGGGCCGC CGTTGACGCA CGTGAAGCAC CGGCTCCTCG
 2551 AGGACTGACA CGTGCTACGA GATTCGATT CCACCGCCG CTTCTATGAA
 TCCTGACTGT GCACGATGCT CTAAAGCTAA GGTGGCGGCG GAAGATACTT
 2601 AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA TGATCCTCCA
 TCCAACCGA AGCCTTAGCA AAAGGCCCTG CGGCCGACCT ACTAGGAGGT
 2651 GCGCGGGGAT CTCATGCTGG AGTTCTTCGC CCACCCCAAC TTGTTTATTG
 CGCGCCCCCTA GAGTACGACC TCAAGAAGCG GGTGGGGTTG AACAAATAAC
 2701 CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT
 GTCGAATATT ACCAATGTT ATTTCGTTAT CGTAGTGTAA AAAGTGTAA
 2751 AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA
 TTTCGTA AAAA AAAGTACGAT AAGATCAACA CAAACAGGT TTGAGTAGTT
 2801 TGTATCTTAT CATGTCTGTA TACCGTCGAC CTCTAGCTAG AGCTGGCGT
 ACATAGAATA GTACAGACAT ATGGCAGCTG GAGATCGATC TCGAACCGCA
 2851 AATCATGGTC ATAGCTGTT CCTGTTGTGAA ATTGTTATCC GCTCACAAATT
 TTAGTACCAAG TATCGACAAA GGACACACTT TAACAATAGG CGAGTGTAA
 2901 CCACACAAACA TACGAGCCGG AAGCATAAAG TGTAAAGCTT GGGGTGCCTA
 GGTGTGTTGT ATGCTCGCC TTGCTATTT ACATTTCGGA CCCCACGGAT
 2951 ATGAGTGAGC TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTCC
 TACTCACTCG ATTGAGTGTAA TAAACGCAA CGCGAGTGTAC GGGCGAAAGG
 3001 AGTCGGGAAA CCTGTCGTG CAGCTGCATT AATGAATCGG CCAACGCCG
 TCAGCCCTTT GGACAGCACG GTCGACGTAA TTACTTAGCC GGTTGCGCGC
 3051 GGGAGAGGCG GTTTGCCTAT TGGGCGCTCT TCCGCTTCC CGCTCACTGA
 CCCTCTCCGC CAAACGCATA ACCCCGGAGA AGGCGAAGGA GCGAGTGA
 3101 CTCGCTCGC TCGGTCGTT GGCTGCGGCG AGCGGTATCA GCTCACTCAA
 GAGCGACGCG AGCCAGCAAG CCGACGCCGC TCGCCATAGT CGAGTGTAGTT
 3151 AGGCGGTAAT ACGGTTATCC ACAGAACGAG GGGATAAACGC AGGAAAGAAC
 TCCGCCATT A TGCCAATAGG TGTCTTAGTC CCCTATTGCG TCCCTTCTTG
 3201 ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCCGCTT
 TACACTCGTT TTCCGGTGT TTTCCGGTCC TTGGCATTTT TCCGGCGCAA
 3251 GCTGGCGTTT TTCCCATAGGC TCCGGCCCCC TGACGAGCAT CACAAAAATC
 CGACCGCAAA AAGGTATCCG AGGCGGGGGG ACTGCTCGTA GTGTTTTAG
 3301 GACGCTCAAG TCAGAGGTGG CGAAACCGA CAGGACTATA AAGATACCAAG
 CTGCGAGTTTC AGTCTCCACC GCTTGGGCT GTCCTGATAT TTCTATGGTC
 3351 GCGTTTCCC CTGGAAAGCTC CCTCGTGC G TCTCTGTT CGACCCCTGCC
 CGAAAGGGGG GACCTTCGAG GGAGCACGCC AGAGGACAAG GCTGGGACGG
 3401 GCTTACCGGA TACCTGTCG CCTTCTCCC TTGCGGAAGC GTGGCGCTT
 CGAATGGCCT ATGGACAGGC GGAAAGAGGG AAGCCCTTCG CACCGCGAAA
 3451 CTCATGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC
 GAGTTACGAG TGCGACATCC ATAGAGTCAA GCCACATCCA GCAAGCGAGG
 3501 AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT
 TTCGACCGA CACACGTGCT TGGGGGGCAA GTCGGGCTGG CGACGCGGAA
 3551 ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAAGACAC GACTTATCGC
 TAGGCCATTG ATAGCAGAAC TCAGGTTGGG CCATTCTGTG CTGAATAGCG

FIG.32D

66/76

3601 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC
 GTGACCGTCG TCGGTGACCA TTGTCCTAAT CGTCTCGCTC CATACATCCG
 3651 GGTGCTACAG AGTTCTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG
 CCACGATGTC TCAAGAACCT CACCACCGGA TTGATGCCGA TGTGATCTTC
 3701 GACAGTATTG GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA
 CTGTCATAAA CCATAGACGC GAGACGACTT CGGTCAATGG AAGCCTTTT
 3751 GAGTTGGTAG CTCTTGATCC GGCAAAACAAA CCACCGCTGG TAGCGGTGGT
 CTCAACCATC GAGAACTAGG CCGTTTGTGTT GGTGGCGACC ATCGCCACCA
 3801 TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTAAGA
 AAAAAACAAA CGTTCGTCGT CTAATGCGCG TCTTTTTTTC CTAGAGTTCT
 3851 AGATCCTTTG ATCTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAAC
 TCTAGGAAAC TAGAAAAGAT GCCCCAGACT GCGAGTCACC TTGCTTTGA
 3901 CACGTTAAGG GATTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG
 GTGCAATTCC CTAAAACCAAG TACTCTAATA GTTTTCTCTA GAAGTGGATC
 3951 ATCCTTTAA ATTAAAAATG AAGTTTTAAA TCAATCTAA GTATATATGA
 TAGGAAAATT TAATTTTAC TTCAAAATT AGTTAGATT CATATATACT
 4001 GTAAAATTTGG TCTGACAGTT ACCAATGCTT AATCAGTGG AGCCTATCT
 CATTGAAACC AGACTGTCAA TGGTTACGAA TTAGTCACTC CGTGGATAGA
 4051 CAGCGATCTG TCTATTCGT TCATCCATAG TTGCTGACT CCCCCTCGTG
 GTCGCTAGAC AGATAAAAGCA AGTAGGTATC AACGGACTGA GGGGCAGCAC
 4101 TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT
 ATCTATTGAT GCTATGCCCT CCCGAATGGT AGACCGGGGT CACGACGTTA
 4151 GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAAC
 CTATGGCGCT CTGGGTGCGA GTGGCCGAGG TCTAAATAGT CGTTATTG
 4201 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCTGCAAC TTTATCCGCC
 TCGGTGGGCC TTCCCGGCTC GCGTCTTCAC CAGGACGTTG AAATAGGCC
 4251 TCCATCCAGT CTATTAATTG TTGCGGGAA GCTAGAGTAA GTAGTTGCC
 AGGTAGGTCA GATAATTAAC AACGGCCCTT CGATCTCATT CATCAAGCGG
 4301 AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC ATCGTGGTGT
 TCAATTATCA AACCGTTCG AACAAACGGTA ACGATGTCGG TAGCACCACA
 4351 CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTTC CCAACGATCA
 GTGCGAGCAG CAAACCATAAC CGAAGTAAGT CGAGGCCAAG GGTTGCTAGT
 4401 AGGGCAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT
 TCCGCTCAAT GTACTAGGGG GTACAAACACG TTTTTGCC AATCGAGGAA
 4451 CGGTCCCTCCG ATCGTTGTC GAAGTAAGTT GGCCGCGAGT TTATCACTCA
 GCCAGGAGGC TAGCAACAGT CTTCATTCAA CCGCGCGTAC AATAGTGAGT
 4501 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA
 ACCAATACCG TCGTGACGTA TTAAGAGAAT GACAGTACGG TAGGCATTCT
 4551 TGCTTTCTG TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG
 ACGAAAAGAC ACTGACCACT CATGAGTTGG TTCAAGTAAGA CTCTTATCAC
 4601 TATGCGGCCGA CCGAGTTGCT CTTGCCCGC GTCAATACGG GATAATACCG
 ATACGCCGCT GGCTCAACGA GAACGGGCCG CAGTTATGCC CTATTATGGC
 4651 CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTCG
 GCGGTGTATC GTCTGAAAT TTTCACGAGT AGTAACCTTT TGCAAGAAC
 4701 GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA
 CCCGCTTTG AGAGTTCTA GAATGGCGAC AACTCTAGGT CAAGCTACAT
 4751 ACCCACTCGT GCACCCAAC GATCTTCAGC ATCTTTTACT TTCACCAAGCG
 TGGGTGAGCA CGTGGGTGTA CTAGAAGTCG TAGAAAATGA AAGTGGTCGC

FIG.32E

67/76

4801 TTTCTGGGTG AGCAAAAC A GGAAGGCAAA ATGCCGCAAA AAAGGGAATA
AAAGACCCAC TCGTTTTGT CCTTCGTTT TACGGCGTTT TTTCCCTTAT
4851 AGGGCGACAC GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA
TCCCGCTGTG CCTTTACAAC TTATGAGTAT GAGAAGGAAA AAGTTATAAT
4901 TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTTGAAT
AACTTCGTAA ATAGTCCCAA TAACAGAGTA CTCGCCTATG TATAAACTTA
4951 GTATTTAGAA AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA
CATAAATCTT TTTATTTGTT TATCCCCAAG GCGCGTGTAA AGGGGCTTTT
5001 GTGCCACCTG ACGTC
CACGGTGGAC TGCAG

FIG.32F

68/76

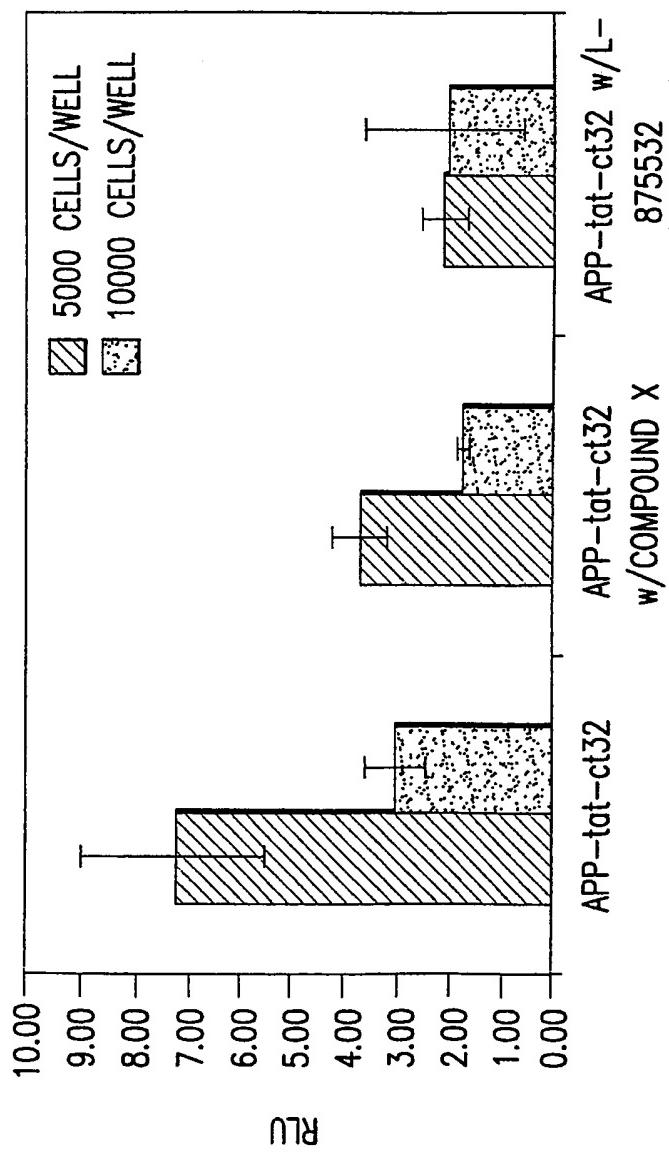
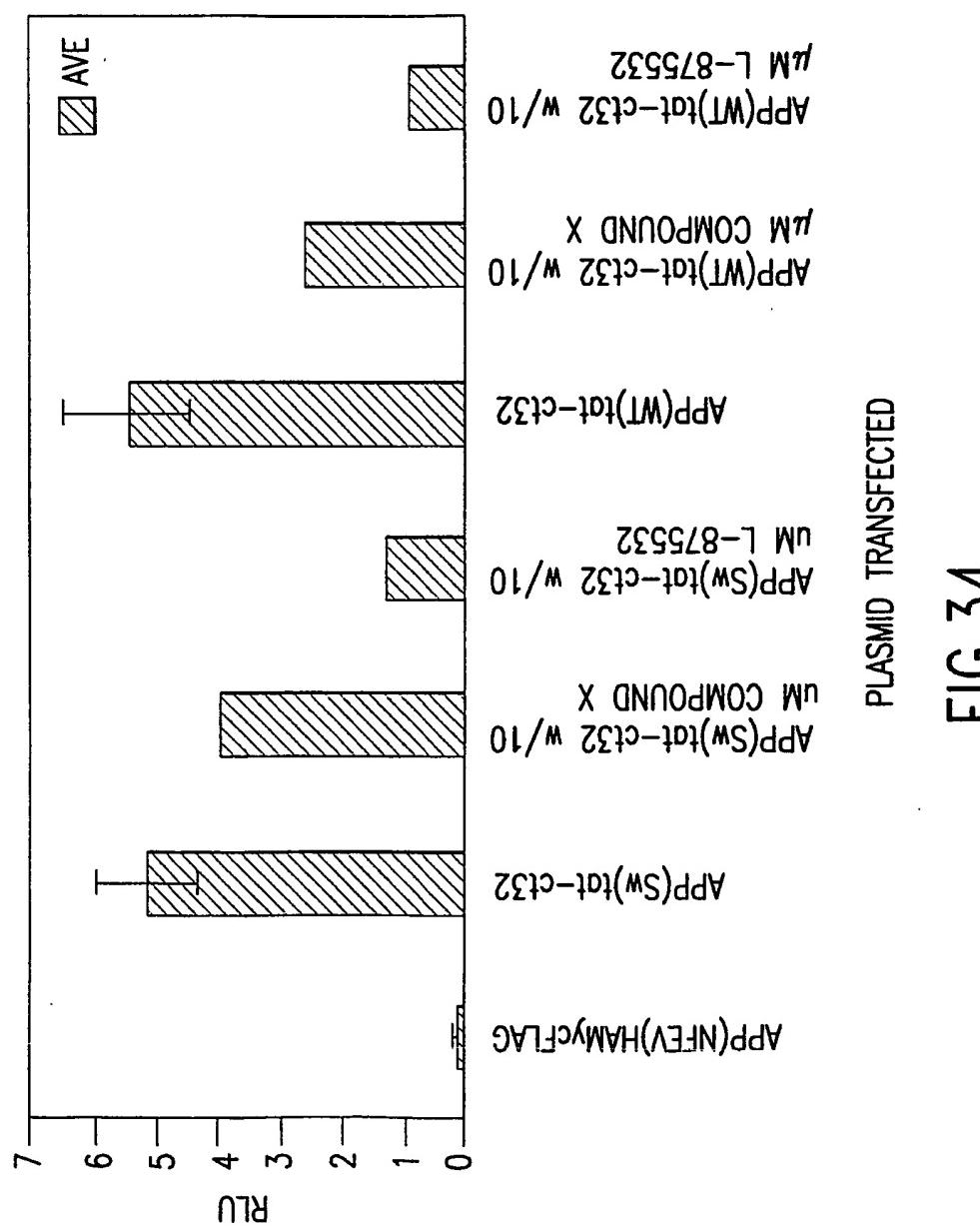


FIG. 33

69/76



70/76

DNA sequence of APP(1-651)NFEV, TATexon1(M1L) APP (664-695)
(SEQ ID NO: 29)

1 ATGCTGCCG GTTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCGCAAGC AGTGCAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
351 CTTAGTTGGT GAGTTATAA GTGATGCCCT TCTCGTTCCCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCAGTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCAGGAGCAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACCAACC AGCATTGCCA CCACCACAC CACCACACAC GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCCCT CCTGCCTCGTC ACGTGTCAA
1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGACAGA CAGCACACCC
1301 TAAAGCATTG CGAGCATGTG CGCATGGTGG ATCCCAAGAA AGCCGCTCAG

FIG.35A

71/76

1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCA GTGGCC GAGGAGATT
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAACTTGA AGTGGAAATT
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTGC
1851 AGAAGATGTG GGTTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGCG
1901 GTGTTGTAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAAATT
2001 TGCAGATATC CTGGAGCCAG TAGATCCTAG ACTAGAGCCC TGGAAGCATC
2051 CAGGAAGTCA GCCTAAAAC GCTTGTACCA ATTGCTATTG TAAAAAGTGT
2101 TGCTTTCATT GCCAAGTTG TTTCATGACA AAAGCCTTAG GCATCTCCTA
2151 TGGCAGGAAG AAGCGGAGAC AGCGACGAAG AGCTCATCAG AACAGTCAGA
2201 CTCATCAAGC TTCTCTATCA AAGCAGAGGA TATCCAGCAC AGTGGCGGCC
2251 GCAGACGCCG CTGTCACCCC AGAGGAGCGC CACCTGTCCA AGATGCAGCA
2301 GAACGGCTAC GAAAATCAA CCTACAAGTT CTTTGAGCAG ATGCAGAACT
2351 AG

FIG.35B

72/76

(SEQ ID NO: 30)

Amino acid sequence of APP(1-651)NFEV, K612V-TATexonI(M1L) APP (664-695)

m1pglal111aawtaralevptdgnag11aepqiamfcgr1nmhm1vqngkwdsdpsgtktcidtkegilqycq
evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisdallvpdkckf1hqermvdvceth1h
whtvaketcsekstn1hdym11pcgidkfrgvefvccplaeesdnvsadaeeddsdvwggadtdyadgs
1
edkvvevaeeeevaeveeeeadddedgedveeeeapeyeaterttsiatttttesveevrvptaastpd
avdkyletpgdenehahfqkakerleakhrermsqvmoreeweaerqaknlpkadkkaviqhfkqekvesleqe
aanerqq1vethmarveam1ndrrrlaleniyitalqavpprprhvf1mkkkyvraeqkdrqht1khfehvrmvd
pkkaaqirsqvmth1rviyermnqs1s1lynvpavaeeiqdevde11qkeqnyssdv1anmisseprisygnda1
mps1tetkttvel1pvngefs1dd1qpwhsfadsvpantenevevpdarpaadrg1ttrpgsg1tnikteeisev
2 3 4 5
nfeverfrhdsgyevhhqk1vffaedvgsnkaiig1mvggvviatvivit1vm1kkk1gtelgstspwwwns
6
ad1lepvdp1lepwkpgsqpktaactncycckccfhcqvcfm1kalgi1sygrkkrrrrrahqnsqthqas1skq
7 8
r1sstvaaadaavtpeerh1skmqngyenptykffeqmgn

FIG.36

73/76

DNA sequence of APP(1-651)NFEV, GAL4VP16(de1Met) APP (664-695)
(SEQ ID NO: 31)

1 ATGCTGCCG GTTGGCACT GCTCCTGCTG GCCGCCTGGA CGGCTCGGGC
51 GCTGGAGGTA CCCACTGATG GTAATGCTGG CCTGCTGGCT GAACCCAGA
101 TTGCCATGTT CTGTGGCAGA CTGAACATGC ACATGAATGT CCAGAACATGGG
151 AAGTGGGATT CAGATCCATC AGGGACCAAA ACCTGCATTG ATACCAAGGA
201 AGGCATCCTG CAGTATTGCC AAGAAGTCTA CCCTGAACCTG CAGATCACCA
251 ATGTGGTAGA AGCCAACCAA CCAGTGACCA TCCAGAACTG GTGCAAGCGG
301 GGCGCGAACG AGTGAAGAC CCATCCCCAC TTTGTGATTC CCTACCGCTG
351 CTTAGTTGGT GAGTTTATAA GTGATGCCCT TCTCGTTCT GACAAGTGCA
401 AATTCTTACA CCAGGAGAGG ATGGATGTTT GCGAAACTCA TCTTCAGTGG
451 CACACCGTCG CCAAAGAGAC ATGCAGTGAG AAGAGTACCA ACTTGATGA
501 CTACGGCATG TTGCTGCCCT GCGGAATTGA CAAGTTCCGA GGGGTAGAGT
551 TTGTGTGTTG CCCACTGGCT GAAGAAAGTG ACAATGTGGA TTCTGCTGAT
601 GCGGAGGAGG ATGACTCGGA TGTCTGGTGG GGCAGGACAG ACACAGACTA
651 TGCAGATGGG AGTGAAGACA AAGTAGTAGA AGTAGCAGAG GAGGAAGAAG
701 TGGCTGAGGT GGAAGAAGAA GAAGCCGATG ATGACGAGGA CGATGAGGAT
751 GGTGATGAGG TAGAGGAAGA GGCTGAGGAA CCCTACGAAG AAGCCACAGA
801 GAGAACCAACC AGCATTGCCA CCACCACCA CACCACCA GAGTCTGTGG
851 AAGAGGTGGT TCGAGTTCCCT ACAACAGCAG CCAGTACCCC TGATGCCGTT
901 GACAAGTATC TCGAGACACC TGGGGATGAG AATGAACATG CCCATTTCCA
951 GAAAGCCAAA GAGAGGCTTG AGGCCAAGCA CCGAGAGAGA ATGTCCAGG
1001 TCATGAGAGA ATGGGAAGAG GCAGAACGTC AAGCAAAGAA CTTGCCTAAA
1051 GCTGATAAGA AGGCAGTTAT CCAGCATTTC CAGGAGAAAG TGGAATCTTT
1101 GGAACAGGAA GCAGCCAACG AGAGACAGCA GCTGGTGGAG ACACACATGG
1151 CCAGAGTGGA AGCCATGCTC AATGACCGCC GCCGCCTGGC CCTGGAGAAC
1201 TACATCACCG CTCTGCAGGC TGTTCTCCT CGGCCTCGTC ACGTGTCAA

FIG.37A

74/76

1251 TATGCTAAAG AAGTATGTCC GCGCAGAACAA GAAGGACAGA CAGCACACCC
1301 TAAAGCATT CGAGCATGTG CGCATGGTGG ATCCAAGAA AGCCGCTCAG
1351 ATCCGGTCCC AGGTTATGAC ACACCTCCGT GTGATTATG AGCGCATGAA
1401 TCAGTCTCTC TCCCTGCTCT ACAACGTGCC TGCAAGTGGCC GAGGAGATT
1451 AGGATGAAGT TGATGAGCTG CTTCAGAAAG AGCAAAACTA TTCAGATGAC
1501 GTCTTGGCCA ACATGATTAG TGAACCAAGG ATCAGTTACG GAAACGATGC
1551 TCTCATGCCA TCTTGACCG AAACGAAAAC CACCGTGGAG CTCCCTCCG
1601 TGAATGGAGA GTTCAGCCTG GACGATCTCC AGCCGTGGCA TTCTTTGGG
1651 GCTGACTCTG TGCCAGCAA CACAGAAAAC GAAGTTGAGC CTGTTGATGC
1701 CCGCCCTGCT GCCGACCGAG GACTGACCAC TCGACCAGGT TCTGGGTTGA
1751 CAAATATCAA GACGGAGGAG ATCTCTGAAG TGAACTTTGA AGTGGAAATT
1801 CGACATGACT CAGGATATGA AGTTCATCAT CAAAAATTGG TGTTCTTTGC
1851 AGAAAGATGTG GGTCAAACA AAGGTGCAAT CATTGGACTC ATGGTGGGCG
1901 GTGTTGTAT AGCGACAGTG ATCGTCATCA CCTTGGTGAT GCTGAAGAAG
1951 AAAAGCTTG GTACCGAGCT CGGATCCACT AGTCCAGTGT GGTGGAATT
2001 TGCAGATATC AAGCTACTGT CTTCTATCGA ACAAGCATGC GATATTGCG
2051 GACTTAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT
2101 CTGAAGAACAA ACTGGGAGTG TCGCTACTCT CCCAAAACCA AAAGGTCTCC
2151 GCTGACTAGG GCACATCTGA CAGAAGTGGG ATCAAGGCTA GAAAGACTGG
2201 AACAGCTATT TCTACTGATT TTTCCTCGAG AAGACCTTGA CATGATTTG
2251 AAAATGGATT CTTTACAGGA TATAAAAGCA TTGTTAACAG GATTATTTGT
2301 ACAAGATAAT GTGAATAAG ATGCCGTAC AGATAGATTG GCTTCAGTGG
2351 AGACTGATAT GCCTCTAACAA TTGAGACAGC ATAGAATAAG TGCGACATCA
2401 TCATCGGAAG AGAGTAGTAA CAAAGGTCAA AGACAGTTGA CTGTATCGGG
2451 AATTCCCGGG GATCTGGCCC CCCCCGACCGA TGTCAGCCTG GGGGACGAGC
2501 TCCACTTAGA CGGCGAGGAC GTGGCGATGG CGCATGCCGA CGCGCTAGAC
2551 GATTTCGATC TGGACATGTT GGGGGACGGG GATTCCCCGG GTCCGGGATT

FIG.37B

75/76

2601 TACCCCCCAC GACTCCGCC CCTACGGCGC TCTGGATATG GCCGACTTCG
2651 AGTTTGAGCA GATGTTTACC GATGCCCTTG GAATTGACGA GTACGGTGGG
2701 GATATCCAGC ACAGTGGCGG CCGCGACGCC GCTGTCACCC CAGAGGAGCG
2751 CCACCTGTCC AAGATGCAGC AGAACGGCTA CGAAAATCCA ACCTACAAGT
2801 TCTTTGAGCA GATGCAGAAC TAG

FIG.37C

76/76

(SEQ ID NO: 32)

Amino acid sequence of APP(1-651)NFEV, GAL4-VP16(deM1) APP (664-695)

m1pglal111aawtaralevptdgnagllaepqiamfcgr1nmhmnnvqngkwdspsgtktcidtkegilqycg
evypelqitnvveanqpvtiqnwckrgrkqckthphfvipyrclvgefisda1vpdkckf1hqermdvceth1h
whtvaketcsekstn1hdym11pcgidkfrgvefvccplaeesdnvdssadaeeddsdvwggadtdyadgs
edkvvevaeeeevaeveeeeadddedgedeveeeaeypyeeaterttsiat1ttttesveevrvptaastpd
avdky1etpgdenehahfqkakerleakhrermsqvmreweeaerqakn1pkadkkaviqhfkqekvesleqe
aanerqq1vethmarveam1ndrrrlalenitalqavpprprhvfnm1kkyvraeqkdrqht1khfehvrmd
pkkaaqirsqvmtlrviyermnqs1s1lynvpavaeei1qdevdellqkeqnyssddvlanmisepri sygndal
mps1tetkttvellpvngefs1dd1qpwhsfadsvpantenevepvdarpaadrg1ttrpgsg1tnikteeisev
nfeveyfrhdsgyevhhqk1vffaedvgsnkaiig1mvgyyiatvivit1vm1kkkk1gtelgstspwwwns
ad1k11ssieqacd1crlkk1kcskekpkcakc1knncryspktkrsp1trah1tevesrlerleq1f11ifpred1d
milkm1ds1qdikalltq1fvqd1vnkdavtdrlasvetdmpl1rqhrisatssseessnkqgrq1tvsgipgd1app
tdvs1qde1h1dqedvamahad1ddf1dm1qdqdspqgqftphdsapya1dmadf1fegmftdalqidey
ggdiqhs1gaaadaavtpeerh1skmqqngyenptykffeqmgn

FIG.38

SEQUENCE LISTING

<110> Merck & Co., Inc.
Espeseth, Amy S.
Ferrer, Marc
Flores, Osvaldo A.
Hazuda, Daria J.
Inglese, James
Miller, Michael D.
Register, Bruce
Shi, Xiao-Ping
Simon, Adam J.
Zuck, Paul D.

<120> ASSAYS TO MONITOR AMYLOID PRECURSOR
PROTEIN PROCESSING

<130> 21040-PCT

<150> 60/360,274
<151> 2002-02-27

<160> 64

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 2352
<212> DNA
<213> fusion protein - human

<400> 1

atgctggccg gtttggcaact gtcctgtcgt gcccgcgttga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgcctatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggacaaa 180
acctgcattt ataccaagga aggcatctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcgaagc agtgcacagac ccatccccac tttgtgatcc cttaccgctg cttagtttgt 360
gagtttataa gtgatgcctc ttcgttccat gacaatgtca aatcttataa ccaggagagg 420
atggatgttt gcgaaactca tcgttactgg cacacccgtcg ccaaagagac atgcagttag 480
aagagtacca acttgcatttgc ctacggcatg ttgctgcctt gcggaatttga caagttccga 540
ggggtagatgtt ttgtgtgttgc cccactggctt gaagaaatgtt acaatgttggta ttctgtgtat 600
gcggaggagg atgactcgga tgcgtgttgc ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagttagtata agtagcagag gaggaagaag tggctgtggat ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgtatggg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacca agcatgttgc ccaccacccacca caccacccaca 840
gagttgttgc aagagggtgt tcgatgttctt acaacacgcg ccagtatccc tgatgcgtt 900
gacaagtatc tcgagacacc tggggatgtt aatgaatcatg cccatatttca gaaagccaaa 960
gagaggcttgg aggccaagca ccggagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgtataaga aggcagtttat ccagcatttc 1080
caggagaaatggaaatctt ggaacaggaa gcagccaaacg agagacagca gctgggtggag 1140
acacacatgg ccagagtggaa agccatgttc aatgaccgc gcccgcgttgc cctggagaac 1200
tacatccatgg ctctgtcgaggc tgggttctt cggccgttc acgtgttcaa tatgtctaaag 1260
aagtatgtcc ggcgacaaaca gaaggacaga cagcacaccc taaaggattt cggatgttgc 1320
cgcatggtgg atcccaagaa agccgttcag atccgttccca aggttatgtac acacccgtt 1380
gtgatttatg agcgcatttgc tcagtcttc tccctgtctt acaacgttgc tgcatgttgc 1440
gaggagatcc aggtgttgc tgcgtgttgc cttcgttgc agcaaaaacta ttcatgttgc 1500
gtcttggcca acatgatttgc tgaaccaagg atcaggatgttgc gaaacgttgc tctcatgttgc 1560
tctttgttgc aaacaaaaac caccgtggag cttccatccca taaaggatgttgc gttcgttgc 1620
gacgatctcc agccgtggca ttcttttggg gctgtacttgc tgccagccaa cacagaaaaac 1680
gaagttgttgc ctgttgc cccgcgttgc gccgaccac gactgaccac tcgaccatgttgc 1740

tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctaga tgcagaattc 1800
 cgacatgact caggatatga agttcatcat caagtattgg tgttcttgc agaagatgtg 1860
 ggttcaaaaaca aaggtaaat cattggactc atggggcg gtgttgcat agcgacagtg 1920
 atcgatcatca ctttgtatg gctgaagaag aaaaagcttg gtaccgagct cgatccact 1980
 agtccagtgt ggttgaattc tgcatgatata ctggagccag tagatcctag actagagccc 2040
 tggaaagcatc caggaagtca gcctaaaact gctttagcca attgtcttgg taaaaagtgt 2100
 tgcttcatt gccaagttt tttcatgaca aaaggcttag gcatctccta tggcaggaag 2160
 aagcggagac agcgcacgaa agctcatcag aacagtccaga ctcatcaagc ttctcttatca 2220
 aagcagagga tatccagcac agtggccggc gcagacgccc ctgtcaccccc agaggagcgc 2280
 cacctgtcca agatgcagca gaacggctac gaaaatccaa cttacaagtt ctttgagcag 2340
 atgcagaact ag 2352

<210> 2
 <211> 783
 <212> PRT
 <213> fusion protein - human

<400> 2
 Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1 5 10 15
 Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
 20 25 30
 Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
 35 40 45
 Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
 50 55 60
 Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
 65 70 75 80
 Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
 85 90 95
 Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
 100 105 110
 Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
 115 120 125
 Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
 130 135 140
 Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
 145 150 155 160
 Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
 165 170 175
 Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
 180 185 190
 Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
 195 200 205
 Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
 210 215 220
 Val Val Glu Val Ala Glu Glu Glu Val Ala Glu Val Glu Glu Glu
 225 230 235 240
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
 245 250 255
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Ser Ile
 260 265 270
 Ala Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
 275 280 285
 Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
 290 295 300
 Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
 305 310 315 320
 Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
 325 330 335
 Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
 340 345 350

Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
 355 360 365
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
 370 375 380
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
 385 390 395 400
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
 405 410 415
 Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
 420 425 430
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
 435 440 445
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
 450 455 460
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
 465 470 475 480
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
 485 490 495
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
 500 505 510
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
 515 520 525
 Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
 530 535 540
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
 545 550 555 560
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
 565 570 575
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
 580 585 590
 Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
 595 600 605
 His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
 610 615 620
 Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
 625 630 635 640
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
 645 650 655
 Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu
 660 665 670
 Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro
 675 680 685
 Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys
 690 695 700
 Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys
 705 710 715 720
 Lys Arg Arg Gln Arg Arg Arg Ala His Gln Asn Ser Gln Thr His Gln
 725 730 735
 Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Asp
 740 745 750
 Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn
 755 760 765
 Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
 770 775 780

<210> 3
 <211> 2352
 <212> DNA
 <213> fusion protein - human

<400> 3

atgctgcccg gttggcaact gctcctgctg gccgcctgga cggctcgggc gctggaggt 60
 cccactgatg gtaatgctgg cctgctgct gaaccccaga ttgcacatgt ctgtggcaga 120
 ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
 acctgcattt atacccaagg aggcatcctg cagattgccc aagaagtcta ccctgaactg 240
 cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
 ggccgcaagc agtgcaagac ccataccccac tttgtgattc cctaccgctg cttagttgg 360
 gagtttataa gtgatgccct ttcgttct gacaagtgc aattcttaca ccaggagagg 420
 atggatgttt gcgaaactca ttttcaactgg cacaccgtcg ccaaagagac atgcagttag 480
 aagagtacca acttgcattga ctacggcatg ttgcgtccct gccgaaattga caagttccga 540
 ggggttagagt ttgtgtgtt cccactgct gaagaaaagt acaatgtgga ttctgctgat 600
 gcggaggagg atgatcgga tgcgttggg ggcggagcag acacagacta tgcatgggg 660
 agtgaagaca aagtagtaga agtagcagag gaggagaagatggctgaggat ggaagaagaa 720
 gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
 ccctacgaag aagccacaga gagaaccacc accattgcca ccaccaccac caccaccaca 840
 gagtctgtgg aagaggttgt tcgagtccct acaacacgac ccagtacccc tgatgccgtt 900
 gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
 gagaggctt agggcaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
 gcagaacgtc aagcaaaagaa cttgccttaa gctgataaga aggcatgtt ccagcatttcc 1080
 caggagaaaatggaaatctttt ggaacagaaaa gcaaggaaac agagacagaca gctggggag 1140
 acacacatgg ccagagtggaa agccatgctc aatgaccgccc gccgcctggc cctggagaac 1200
 tacatcaccg ctctgcaggc ttttcttctt cggcctcgac acgtgttcaa tatgctaaag 1260
 aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
 cgcatggtgg atcccaagaa agccgcttag atccggtccc aggttatgac acacccctgt 1380
 gtgattttatg agcgcatgaa tcagtccttc tccctgtct acaacacgtggc tgcatggcc 1440
 gaggagattt aggtatggat tgatgagatg cttcagaaaa agcaaaaacta ttcatgtgac 1500
 gtcttggcca acatgatggat tgaaccaagg atcgttacg gaaacatgtc tctcatgcca 1560
 tctttgaccg aaacaaaaac caccgtggag ctcccttcccg tgaatgggaa gttcagcctg 1620
 gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaaac 1680
 gaagttgagc ctgttcatgac ccgcctgtgt gcccggcggag gactgaccac tcgaccagg 1740
 tctgggttga caaatatcaa gacggaggag atctctgaaatg tgaagatggg tgcatggattc 1800
 cgacatgact caggatgatg agttcatcat caagtattgg ttgttcttgc agaagatgtg 1860
 gttcaaca aaggtgcaat cattggactt atggggggcg gtgttgcattt acgcacagtg 1920
 atcgatcatca ctttggatgatg gctgaaagaaatggggctt gtaccggatc cggatccact 1980
 agtccatgtt ggttggaaatc tgcatgatcatc ctggagccag tagatcttag actagagccc 2040
 tggaaagcatc caggaagtca gcctaaaaact gcttgcacca attgttattt taaaaagtgt 2100
 tgctttcatt gccaagttt tttcatgaca aaaggcttag gcatctccta tggcaggaag 2160
 aagcggagac agcagacgaag agtcatcatag aacagtccaga ctcatcaagc ttctctatca 2220
 aagcagagga tatccagcac agtggcgcc gcagacgccc ctgtcaccctt agaggagcgc 2280
 caccgttcca agatgcagca gaacggctac gaaaatccaa cttacaagttt ctttgcgg 2340
 atgcagaact ag 2352

<210> 4

<211> 783

<212> PRT

<213> fusion protein - human

<400> 4

Met	Leu	Pro	Gly	Leu	Ala	Leu	Leu	Leu	Ala	Ala	Trp	Thr	Ala	Arg	
1				5				10			15				
Ala	Leu	Glu	Val	Pro	Thr	Asp	Gly	Asn	Ala	Gly	Leu	Leu	Ala	Glu	Pro
				20				25			30				
Gln	Ile	Ala	Met	Phe	Cys	Gly	Arg	Leu	Asn	Met	His	Met	Asn	Val	Gln
				35				40			45				
Asn	Gly	Lys	Trp	Asp	Ser	Asp	Pro	Ser	Gly	Thr	Lys	Thr	Cys	Ile	Asp
			50						55		60				
Thr	Lys	Glu	Gly	Ile	Leu	Gln	Tyr	Cys	Gln	Glu	Val	Tyr	Pro	Glu	Leu
				65				70		75			80		
Gln	Ile	Thr	Asn	Val	Val	Glu	Ala	Asn	Gln	Pro	Val	Thr	Ile	Gln	Asn
				85				90			95				
Trp	Cys	Lys	Arg	Gly	Arg	Lys	Gln	Cys	Lys	Thr	His	Pro	His	Phe	Val
				100				105			110				
Ile	Pro	Tyr	Arg	Cys	Leu	Val	Gly	Glu	Phe	Ile	Ser	Asp	Ala	Leu	Leu

115	120	125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys		
130	135	140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu		
145	150	155
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile		
165	170	175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu		
180	185	190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val		
195	200	205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys		
210	215	220
Val Val Glu Val Ala Glu Glu Glu Val Ala Glu Val Glu Glu Glu		
225	230	235
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu		
245	250	255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile		
260	265	270
Ala Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg		
275	280	285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu		
290	295	300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys		
305	310	315
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg		
325	330	335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp		
340	345	350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu		
355	360	365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala		
370	375	380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn		
385	390	395
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe		
405	410	415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His		
420	425	430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala		
435	440	445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu		
450	455	460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala		
465	470	475
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn		
485	490	495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser		
500	505	510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr		
515	520	525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln		
530	535	540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn		
545	550	555
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr		
565	570	575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser		
580	585	590
Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val		
595	600	605
His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys		

610	615	620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val		
625	630	640
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Leu Gly Thr Glu		
	645	655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu		
	660	670
Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro		
	675	685
Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys		
	690	700
Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys		
	705	720
Lys Arg Arg Gln Arg Arg Arg Ala His Gln Asn Ser Gln Thr His Gln		
	725	735
Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Ala Asp		
	740	750
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn		
	755	765
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn		
	770	780

<210> 5
<211> 2823
<212> DNA
<213> fusi

agtccagtgt ggtggaattc tgcagatatac aagctactgt cttcttatcga acaagcatgc 2040
 gatatttgcc gactaaaaaa gctcaagtgc tccaaagaaa aaccgaagtg cgccaagtgt 2100
 ctgaagaaca actgggagtg tcgctactct cccaaaacca aaaggcttcg gctgactagg 2160
 gcacatctga cagaagtggta atcaaggctta gaaagactgg aacagctatt tctactgatt 2220
 tttcctcgag aagaccttga catgatttg aaaatggatt ctttacagga tataaaagca 2280
 ttgttaacag gattattgt acaagataat gtgaataaag atgccgtcac agatagattg 2340
 gcttcagtgg agactgatat gcctctaaca ttgagacage atagaataag tgcgacatca 2400
 tcatcgaag agagtagtaa caaaggtaa agacagtta ctgtatcggg aattcccccggg 2460
 gatctggccc ccccgaccga tgtcagctg ggggacgagc tccacttaga cggcgaggac 2520
 gtggcgatgg cgcacatggca cgcgctagatc gatttcgatc tggacatgtt gggggacggg 2580
 gattccccgg gtcgggatt taccggggcc gactccggcc cctacggcgc tctggatatg 2640
 gccgacttcg agtttgagca gatgtttacc gatggccctt gatggacgca gtacggtggg 2700
 gatatccagc acagtggcgcc cgcgcacgccc gctgtcaccc cagaggagcg ccacctgtcc 2760
 aagatgcagc agaacggcta cggaaatcca acctacaagt tctttgagca gatgcagaac 2820
 tag

<210> 6

<211> 941

<212> PRT

<213> fusion protein - human

<400> 6

Met	Leu	Pro	Gly	Leu	Ala	Leu	Leu	Leu	Leu	Ala	Ala	Trp	Thr	Ala	Arg
1	5											10		15	
Ala	Leu	Glu	Val	Pro	Thr	Asp	Gly	Asn	Ala	Gly	Leu	Leu	Ala	Glu	Pro
		20							25				30		
Gln	Ile	Ala	Met	Phe	Cys	Gly	Arg	Leu	Asn	Met	His	Met	Asn	Val	Gln
		35						40			45				
Asn	Gly	Lys	Trp	Asp	Ser	Asp	Pro	Ser	Gly	Thr	Lys	Thr	Cys	Ile	Asp
		50						55			60				
Thr	Lys	Glu	Gly	Ile	Leu	Gln	Tyr	Cys	Gln	Glu	Val	Tyr	Pro	Glu	Leu
		65					70			75			80		
Gln	Ile	Thr	Asn	Val	Val	Glu	Ala	Asn	Gln	Pro	Val	Thr	Ile	Gln	Asn
							85			90			95		
Trp	Cys	Lys	Arg	Gly	Arg	Lys	Gln	Cys	Lys	Thr	His	Pro	His	Phe	Val
		100					105				110				
Ile	Pro	Tyr	Arg	Cys	Leu	Val	Gly	Glu	Phe	Ile	Ser	Asp	Ala	Leu	Leu
		115					120				125				
Val	Pro	Asp	Lys	Cys	Lys	Phe	Leu	His	Gln	Glu	Arg	Met	Asp	Val	Cys
		130				135				140					
Glu	Thr	His	Leu	His	Trp	His	Thr	Val	Ala	Lys	Glu	Thr	Cys	Ser	Glu
		145				150				155			160		
Lys	Ser	Thr	Asn	Leu	His	Asp	Tyr	Gly	Met	Leu	Leu	Pro	Cys	Gly	Ile
							165			170			175		
Asp	Lys	Phe	Arg	Gly	Val	Glu	Phe	Val	Cys	Cys	Pro	Leu	Ala	Glu	Glu
		180				185					190				
Ser	Asp	Asn	Val	Asp	Ser	Ala	Asp	Ala	Glu	Glu	Asp	Asp	Ser	Asp	Val
		195				200				205					
Trp	Trp	Gly	Gly	Ala	Asp	Thr	Asp	Tyr	Ala	Asp	Gly	Ser	Glu	Asp	Lys
		210				215				220					
Val	Val	Glu	Val	Ala	Glu	Glu	Glu	Val	Ala	Glu	Val	Glu	Glu		
		225				230				235			240		
Glu	Ala	Asp	Asp	Asp	Glu	Asp	Asp	Glu	Asp	Gly	Asp	Glu	Val	Glu	Glu
							245			250			255		
Glu	Ala	Glu	Glu	Pro	Tyr	Glu	Glu	Ala	Thr	Glu	Arg	Thr	Thr	Ser	Ile
		260				265				270					
Ala	Thr	Thr	Thr	Thr	Thr	Thr	Glu	Ser	Val	Glu	Glu	Val	Val	Val	Arg
		275				280				285					
Val	Pro	Thr	Thr	Ala	Ala	Ser	Thr	Pro	Asp	Ala	Val	Asp	Lys	Tyr	Leu
		290				295				300					
Glu	Thr	Pro	Gly	Asp	Glu	Asn	Glu	His	Ala	His	Phe	Gln	Lys	Ala	Lys
		305				310				315			320		

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
 325 330 335
 Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
 340 345 350
 Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
 355 360 365
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
 370 375 380
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
 385 390 395 400
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
 405 410 415
 Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
 420 425 430
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
 435 440 445
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
 450 455 460
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
 465 470 475 480
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
 485 490 495
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
 500 505 510
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
 515 520 525
 Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
 530 535 540
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
 545 550 555 560
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
 565 570 575
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
 580 585 590
 Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
 595 600 605
 His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
 610 615 620
 Gly Ala Ile Ile Gly Leu Met Val Gly Val Val Ile Ala Thr Val
 625 630 635 640
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Leu Gly Thr Glu
 645 650 655
 Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
 660 665 670
 Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
 675 680 685
 Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
 690 695 700
 Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
 705 710 715 720
 Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
 725 730 735
 Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
 740 745 750
 Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
 755 760 765
 Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
 770 775 780
 Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
 785 790 795 800
 Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
 805 810 815

Gly	Ile	Pro	Gly	Asp	Leu	Ala ^a	Pro	Pro	Thr	Asp	Val	Ser	Leu	Gly	Asp
			820				825					830			
Glu	Leu	His	Leu	Asp	Gly	Glu	Asp	Val	Ala	Met	Ala	His	Ala	Asp	Ala
			835				840					845			
Leu	Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Gly	Asp	Gly	Asp	Ser	Pro	Gly
			850				855					860			
Pro	Gly	Phe	Thr	Pro	His	Asp	Ser	Ala	Pro	Tyr	Gly	Ala	Leu	Asp	Met
			865				870					875			880
Ala	Asp	Phe	Glu	Phe	Glu	Gln	Met	Phe	Thr	Asp	Ala	Leu	Gly	Ile	Asp
			885				890					895			
Glu	Tyr	Gly	Gly	Asp	Ile	Gln	His	Ser	Gly	Ala	Ala	Ala	Asp	Ala	Ala
			900				905					910			
Val	Thr	Pro	Glu	Glu	Arg	His	Leu	Ser	Lys	Met	Gln	Gln	Asn	Gly	Tyr
			915				920					925			
Glu	Asn	Pro	Thr	Tyr	Lys	Phe	Phe	Glu	Gln	Met	Gln	Asn			
			930				935					940			

```

<210> 7
<211> 2823
<212> DNA
<213> fusion protein - human

<400> 7
atgtgcggcc gtttggact gtcctgctg gccgcctgga cggctcgccc gctggaggta 60
cccactgtat gtaatgctgg ctgtctggct gaacccaga ttggcatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattt ataccaagga aggcatctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tcoagaactg gtgcaaggc 300
ggccgcgaagc agtgcaagac ccatcccccc ttgtgttcc cttaccgctg cttagttgt 360
gagtttataa gtgtatccct ttcgttccct gacaagtca aattttaca ccaggagagg 420
atgtatgtt gggaaactca ttcttactgg cacacccgtcg ccaaaagagc atgcgttgag 480
aagagtacca acttgcattgt ctacggcatg ttgctccct gggaaattgtt caagttccg 540
ggggtagagt ttgtgttgg cccactggct gaagaaaatgtt acaatgttggg ttctgtctgt 600
gcccggaggagg atgactcgga tgtctgggtt ggcggagcag acacagacta tgcatgtggg 660
agtgaagaca aagttagttaga agtagcagag gaggaaagaatg tggctgaggat ggaagaagaa 720
gaagccgtat atgacggaggat cgatgggatgtt ggtgtatggg tagaggaaaga ggcttagggaa 780
ccctacacaaag aaggccaccaaa gagaaccaccatc agcattgcctt ccaccacccac 840
gagtctgtgg aagagggttgtt tcgatgttccct acaacacggc ccagtatcccccc 900
gacaagtatc tcgagacacc tggggatgtt aatgaacatg cccatccatc gaaaagccaaa 960
gagaggcttgg aggccaaagca ccgagagagaa atgtcccttggg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaaagaa cttgcctaaa gctgataaga aggcagtat cccagcattt 1080
caggagaaag tggaaatctttt ggaacaggaa gcagccaaacg agagacagca gctgggtggag 1140
acacacatgg ccagagtgaa agccatgtctt aatgaccggcc gcccctgtcc cctggagaaac 1200
tacatcaccc tcctcgaggc tgggttccctt cggccctgtc acgtttcaat tatgtctaaag 1260
aagtatgtcc ggcgcaaaaca aaggaggacaga cagcacaccc taaggatccatcg cgacatgttgg 1320
cgcatggggg atcccaaaagaa agccgtcttag atccggtccc aggttatgac acacccctcg 1380
gtgatattatg agcgcatgaa tcagttcttc tccctgtctt acaacgtgcc tgcatgtggcc 1440
gaggagatcc aggtatgttggatgt tgatgagctg cttcagaaaag agcaaaaacta ttcatgttgc 1500
gtcttggcca acatgattttt gtaaccaagg atcagttacg gaaacgtatc tctcatgcca 1560
tctttgaccc aaacggaaaac caccgtggag ctccctcccg tgaatggaga gttcagctg 1620
gacgtatctcc agccgtggca ttctttttgg gctgactctg tgccagccaa cacaaaaaac 1680
gaagtgtggc ctgttgatgc ccccccgtctt gccgaccggag gactgaccac tcgaccagg 1740
tctgggttgc caaatatcaa gacggaggag atctctgttca gtaagatggaa tgccagaattt 1800
cgacatgact caggatattgt agttcatcat caagtattgg tggatgttgc agaagatgtt 1860
ggttcaaaaca aagggtcaat cattggactc atgggtggcg gtgttgcattt agcgacatgt 1920
atcgatcatca ctttgtgtat gctgaagaaag aaaaagctt gtaaccggactt cggatccact 1980
agtccagttt ggtggaaattt tgcatgttccctt aagctactgt cttctatcgat acaagcatgc 2040
gatatttgc gactttttttt gctcaagtgc tccaaagaaaaa aacccgaatgtt cgccaaatgtt 2100
ctgaagaaca actggggatgt tgctgttccctt cccaaaacca aaaaatgttcc gctgtactgg 2160
gcacatctgtt cagaagtggaa atcaaggctt gaaagacttgg aacacatgtt tctactgtt 2220
ttttccctgtt gaaagacttgc catgttattttt gaaatgttggat ctttacaggatataaaaagca 2280

```

ttgttaacag gattatttgc acaagataat gtgaataaag atgccgtcac agatagattg 2340
 gcttcagtgg agactgatat gcctctaaca ttgagacagc atagaataag tgcgacatca 2400
 tcatcgaaag agagtagtat caaaggtaa agacagtga ctgtatcgaa aattcccgaa 2460
 gatctggccc ccccgaccga tgctagccgt ggggacgagc tccacttaga cggcgaggac 2520
 gtggcgatgg cgcatgccga cgcgctagac gatttcgatc tggacatgtt gggggacggg 2580
 gattccccgg gtcgggatt taccggccac gactccgccc cctacggcgc tctggatatg 2640
 gccgacttcg agtttgagca gatgtttacc gatgcccttga atttgacgca gtacgggtgg 2700
 gatatccagc acagtggcg ccgcgacgccc gctgtcaccc cagaggagcg ccacctgtcc 2760
 aagatgcagc agaacggcta cgaaaatcca acctacaagt tctttgagca gatgcagaac 2820
 tag 2823

<210> 8

<211> 941

<212> PRT

<213> fusion protein - human

<400> 8

Met	Leu	Pro	Gly	Leu	Ala	Leu	Leu	Leu	Ala	Ala	Trp	Thr	Ala	Arg	
1	5						10				15				
Ala	Leu	Glu	Val	Pro	Thr	Asp	Gly	Asn	Ala	Gly	Leu	Leu	Ala	Glu	Pro
		20					25				30				
Gln	Ile	Ala	Met	Phe	Cys	Gly	Arg	Leu	Asn	Met	His	Met	Asn	Val	Gln
	35						40				45				
Asn	Gly	Lys	Trp	Asp	Ser	Asp	Pro	Ser	Gly	Thr	Lys	Thr	Cys	Ile	Asp
	50						55				60				
Thr	Lys	Glu	Gly	Ile	Leu	Gln	Tyr	Cys	Gln	Glu	Val	Tyr	Pro	Glu	Leu
	65				70				75						80
Gln	Ile	Thr	Asn	Val	Val	Glu	Ala	Asn	Gln	Pro	Val	Thr	Ile	Gln	Asn
		85					90							95	
Trp	Cys	Lys	Arg	Gly	Arg	Lys	Gln	Cys	Lys	Thr	His	Pro	His	Phe	Val
	100						105							110	
Ile	Pro	Tyr	Arg	Cys	Leu	Val	Gly	Glu	Phe	Ile	Ser	Asp	Ala	Leu	Leu
	115					120						125			
Val	Pro	Asp	Lys	Cys	Lys	Phe	Leu	His	Gln	Glu	Arg	Met	Asp	Val	Cys
	130					135					140				
Glu	Thr	His	Leu	His	Trp	His	Thr	Val	Ala	Lys	Glu	Thr	Cys	Ser	Glu
	145					150				155					160
Lys	Ser	Thr	Asn	Leu	His	Asp	Tyr	Gly	Met	Leu	Leu	Pro	Cys	Gly	Ile
		165					170							175	
Asp	Lys	Phe	Arg	Gly	Val	Glu	Phe	Val	Cys	Cys	Pro	Leu	Ala	Glu	Glu
	180					185								190	
Ser	Asp	Asn	Val	Asp	Ser	Ala	Asp	Ala	Glu	Glu	Asp	Asp	Ser	Asp	Val
	195					200						205			
Trp	Trp	Gly	Gly	Ala	Asp	Thr	Asp	Tyr	Ala	Asp	Gly	Ser	Glu	Asp	Lys
	210					215					220				
Val	Val	Glu	Val	Ala	Glu	Glu	Glu	Val	Ala	Glu	Val	Glu	Glu		
	225					230				235				240	
Glu	Ala	Asp	Asp	Glu	Asp	Asp	Glu	Asp	Gly	Asp	Glu	Val	Glu	Glu	
	245					250				255					
Glu	Ala	Glu	Glu	Pro	Tyr	Glu	Glu	Ala	Thr	Glu	Arg	Thr	Ser	Ile	
	260					265					270				
Ala	Thr	Thr	Thr	Thr	Thr	Thr	Glu	Ser	Val	Glu	Glu	Val	Val	Arg	
	275					280					285				
Val	Pro	Thr	Thr	Ala	Ala	Ser	Thr	Pro	Asp	Ala	Val	Asp	Lys	Tyr	Leu
	290					295					300				
Glu	Thr	Pro	Gly	Asp	Glu	Asn	Glu	His	Ala	His	Phe	Gln	Lys	Ala	Lys
	305					310				315				320	
Glu	Arg	Leu	Glu	Ala	Lys	His	Arg	Glu	Arg	Met	Ser	Gln	Val	Met	Arg
	325					330						335			
Glu	Trp	Glu	Glu	Ala	Glu	Arg	Gln	Ala	Lys	Asn	Leu	Pro	Lys	Ala	Asp
	340					345						350			
Lys	Lys	Ala	Val	Ile	Gln	His	Phe	Gln	Glu	Lys	Val	Glu	Ser	Leu	Glu

355	360	365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala		
370	375	380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn		
385	390	395
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe		
405	410	415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His		
420	425	430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala		
435	440	445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu		
450	455	460
Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala		
465	470	475
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn		
485	490	495
Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser		
500	505	510
Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr		
515	520	525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln		
530	535	540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn		
545	550	555
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr		
565	570	575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser		
580	585	590
Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val		
595	600	605
His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys		
610	615	620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val		
625	630	635
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Leu Gly Thr Glu		
645	650	655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu		
660	665	670
Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu		
675	680	685
Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn		
690	695	700
Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg		
705	710	715
Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu		
725	730	735
Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met		
740	745	750
Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln		
755	760	765
Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu		
770	775	780
Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser		
785	790	795
Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser		
805	810	815
Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp		
820	825	830
Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala		
835	840	845
Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Ser Pro Gly		

850	855	860
Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr	Gly Ala Leu Asp Met	
865	870	875
Ala Asp Phe Glu Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp		880
885	890	895
Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Ala Asp Ala Ala		
900	905	910
Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr		
915	920	925
Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn		
930	935	940

<210> 9
<211> 2352

<212> DNA

<213> fusion protein - human

<400> 9

atgctgcccg gtttggcact gctcctgctg gcccctgga cggctcgccc gctggaggta 60
cccaactgatg gtaatgctgg cctgctggct gaaccccaga ttgcctatgt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggacaaa 180
acctgcattt ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagg agtgcacagg ccatccccac ttttgtgattt cttaccgcgt cttagtttgt 360
gagttataaa gtgtatccct ttcgttccct gacaagtgc aattcttaca ccaggagagg 420
atggatgtt gcgaaactca ttttcaactgg cacaccgtc ccaaagagac atgcagttag 480
aagagtacca acttgcacatca ctacggcatg ttgtgtccct gccggaaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgtgtat 600
gcggaggagg atgactcgga tgtctgggtt ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagttagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgtatggat ggtgtatggg tagaggaaga ggctgagggaa 780
cccttacaaag aaggccacaga gagaaccacc agcatgttca ccaccaccac caccacca 840
gagtctgtgg aagagggtgt tcgatgttccct acaacacgcg ccagtatccc tgatgcgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttgg aggccaagca ccggagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcaatttat ccagcatttc 1080
caggagaaat tggaaatctt ggaacaggaa gcagccaacg agagacagca gctgggtggag 1140
acacacatgg ccacatggaa agccatgttcc aatgaccgcg gccgcctggc cctggagaac 1200
tacatccatcg ctctgcaggc tggttccctcc cggccatgtc acgtgttcaaa tatgtctaaag 1260
aagtatgtcc ggcgacaaaca aagggacaca cagcacaccc taaggcattt cgagcatgtg 1320
cgcatgttgg atcccaagaa agccgtctcgt atccgttccc agttatgac acacccgtt 1380
gtgattttatg agcgtatgaa tcagtcttc tccctgtct acaacatgttca tgcatgttggcc 1440
gaggagatcc aggtgaagt tgatgagctg cttcagaaag agcaaaaacta ttcatgatgac 1500
gtcttggcca acatgattatg tgaaccaagg atcaggatcgt gaaacatgtc tctcatgcca 1560
tctttgaccc aaacaaaaac caccgtggag ctcccttcccg tgaatggaga gttcaggctg 1620
gacgtatccc agccgtggca ttctttggg gctgtatgt tgccagccaa cacagaaaaac 1680
gaagtgtggc ctgttgcgttcc cccgcctgtt gcccggcgg gactgaccac tcgaccagg 1740
tctgggttgc caaatatcaa gacggggaggat atctcttgc tgaatcttgc tgccatgtt 1800
cgacatgact caggatatgt aatgttccatcat caaaaattgg tggcttgc agaagatgtg 1860
ggttcaaaaca aaggttgcattt cattggactc atgggtggcg gtttgcgttcat agcgttgc 1920
atcgtatca ctttgcgttcat gctgttgcattt aaaaagctt gtaaccgttccgtt cggatccact 1980
agtccatgtt ggttgcattt tgcgttgcattt ctggggccat tagatccttgc actagagccc 2040
tggaaatccatc caggatgttca gcttgcattt gctgttgcattt attgttgcattt taaaaatgtt 2100
tgctttcattt gccaatgttgcattt ttttgcattt gccaatgttgcattt aaaaatgtt 2160
aagcggagac agcgtatgttca aacgttgcattt ctttgcattt ctttgcattt ttttgcattt 2220
aagcggagac aagcgtatgttca aacgttgcattt ctttgcattt ctttgcattt ttttgcattt 2280
caccgttccatc agatgttgcattt gccaatgttgcattt aaaaatgtt 2340
atgcgttgcattt ag 2352

<210> 10
<211> 783

<212> PRT

<213> fusion protein - human

<400> 10

Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1 5 10 15
 Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
 20 25 30
 Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
 35 40 45
 Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
 50 55 60
 Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
 65 70 75 80
 Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
 85 90 95
 Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
 100 105 110
 Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
 115 120 125
 Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
 130 135 140
 Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
 145 150 155 160
 Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
 165 170 175
 Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
 180 185 190
 Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
 195 200 205
 Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
 210 215 220
 Val Val Glu Val Ala Glu Glu Glu Val Ala Glu Val Glu Glu Glu
 225 230 235 240
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
 245 250 255
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
 260 265 270
 Ala Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
 275 280 285
 Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
 290 295 300
 Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
 305 310 315 320
 Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
 325 330 335
 Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
 340 345 350
 Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
 355 360 365
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
 370 375 380
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
 385 390 395 400
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
 405 410 415
 Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
 420 425 430
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
 435 440 445
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
 450 455 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
 465 470 475 480
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
 485 490 495
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
 500 505 510
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
 515 520 525
 Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
 530 535 540
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
 545 550 555 560
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
 565 570 575
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
 580 585 590
 Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
 595 600 605
 His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
 610 615 620
 Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
 625 630 635 640
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Leu Gly Thr Glu
 645 650 655
 Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu
 660 665 670
 Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro
 675 680 685
 Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys
 690 695 700
 Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys
 705 710 715 720
 Lys Arg Arg Gln Arg Arg Ala His Gln Asn Ser Gln Thr His Gln
 725 730 735
 Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Asp
 740 745 750
 Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn
 755 760 765
 Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
 770 775 780

<210> 11
 <211> 2352
 <212> DNA
 <213> fusion protein - human

<400> 11
 atgctggccg gtttggcaact gtcctgtgt gccgcctgga cggctcgccc gctggaggt 60
 cccactgatg gtaatgctgg cctgctggct gaaccccaaa ttgcctatgtt ctgtggcaga 120
 ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaaa 180
 acctgcattt ataccaagga aggcatctg cagtattgcc aagaagtcta ccctgaactg 240
 cagatcacca atgtggtaga agccaaaccaa ccagtgcacca tccagaactg gtgcaagcgg 300
 gggcgcaagg agtgcacaaac ccatccccac ttgtgttattt ccttccgcgtt ctttagtttgtt 360
 gagtttataaa gtgtatggccct tctcggttcc gacaatgtca aatttttaca ccaggagagg 420
 atggatgttt gcgaaaactca tcttcacttg cacaccgtcg cccaaagagac atgcagttag 480
 aagagtacca acttgcattt ctacggcatg ttgtgttccctt gccggaaatttga caagttccga 540
 ggggttagagt ttgtgtgtt cccactggctt gaagaaaatgtt acaatgtggaa ttctgtgttat 600
 gcgaggaggat atgactcgaa tggatggccctt ggcggagcag acacagacta tgccatgggg 660
 agtgaagaca aagttagttaga agttagcagag gaggaaagaatggctgaggtt ggaagaagaa 720
 gaagcccgatg atgacgagga cgatgaggat ggtgtatgggg tagaggaaga ggctggggaa 780
 ccctacgaag aagccacaga gagaaccacc accattggcca ccaccaccac caccaccaca 840

gagtctgtgg aagagggtgg tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
 gacaagtata tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
 gagaggctt aggccaagca cggagagaga atgtcccagg tcatgagaga atgggaagag 1020
 gcagaacgtc aaggaaagaa ctgcctaaa gctgataaga aggcagttat ccagcattc 1080
 caggagaaaag tggaatctt ggaacagggaa gcagccaacg agagacagca gctggggag 1140
 acacacatgg ccagagtggc tacatcaccc ctctgcaggg tggctcttc cggcctcgac acgtgttcaa tatgtctaaag 1260
 aagtatgtcc ggcgagaaca gaaggacaga cagcacaccc taaaggcattt cgagcatgtg 1320
 cgcattgggg atcccaagaa agccgctca gatccggtccc aggttatgac acacccgt 1380
 gtgattttatg agcgcattaa tcagtcttc tccctgtct acaacgtgcc tgcaagtggcc 1440
 gaggagattc aggatgaagt tgatgagctc cttcagaagg acaaaaaacta ttcatgatgac 1500
 gtcttggcca acatgattttg tgaaccaagg atcagttacg gaaacatgtc tctcatgcca 1560
 tctttgaccg aaacgaaaac caccgtggag ctccctcccg tgaatggaga gttcagcctg 1620
 gacgatctcc agccgtggca ttctttggg gctgactctg tgccagccaa cacagaaaac 1680
 gaagttgagc ctgttgatgc ccgcctgtc gcccggcggag gactgaccac tcgaccagg 1740
 tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
 cgacatgact caggatgtg agttcatcat caaaaaattgg tggctttggc agaagatgtg 1860
 ggttcaaaaaca aaggtaatcatttggactc atggggggcg gtgttggcat agcgacagt 1920
 atcgtcatca ctttgttggat gctgaagaag aaaaagcttgcgtaccgagct cggatccact 1980
 agtccagtgt ggttggattc tgcatgatc ctggagccag tagatcctag actagagccc 2040
 tggaaagatc caggaagtca gcctaaaact gcttgcatttca attgtctatttgg taaaaggatgt 2100
 tgctttcatt gccaagtttgc ttcatgaca aaaggcttag gcatttccta tggcaggaag 2160
 aagcggagac agcgcacgaa agctcatcag aacagtccaga ctcatcaagg ttctctatca 2220
 aagcagagga tatccagcac agtggccggc gcagacgccc ctgtcaccggc agaggagcgc 2280
 cacctgtcca agatgcagca gaacggctac gaaaatccaa cctacaagtt ctttgagcag 2340
 atgcagaact ag 2352

<210> 12

<211> 783

<212> PRT

<213> fusion protein - human

<400> 12

Met	Leu	Pro	Gly	Leu	Ala	Leu	Leu	Leu	Ala	Ala	Trp	Thr	Ala	Arg	
1	5							10				15			
Ala	Leu	Glu	Val	Pro	Thr	Asp	Gly	Asn	Ala	Gly	Leu	Leu	Ala	Glu	Pro
								20			25		30		
Gln	Ile	Ala	Met	Phe	Cys	Gly	Arg	Leu	Asn	Met	His	Met	Asn	Val	Gln
								35			40		45		
Asn	Gly	Lys	Trp	Asp	Ser	Asp	Pro	Ser	Gly	Thr	Lys	Thr	Cys	Ile	Asp
								50			55		60		
Thr	Lys	Glu	Gly	Ile	Leu	Gln	Tyr	Cys	Gln	Glu	Val	Tyr	Pro	Glu	Leu
								65			70		75		80
Gln	Ile	Thr	Asn	Val	Val	Glu	Ala	Asn	Gln	Pro	Val	Thr	Ile	Gln	Asn
								85			90		95		
Trp	Cys	Lys	Arg	Gly	Arg	Lys	Gln	Cys	Lys	Thr	His	Pro	His	Phe	Val
								100			105		110		
Ile	Pro	Tyr	Arg	Cys	Leu	Val	Gly	Glu	Phe	Ile	Ser	Asp	Ala	Leu	Leu
								115			120		125		
Val	Pro	Asp	Lys	Cys	Lys	Phe	Leu	His	Gln	Glu	Arg	Met	Asp	Val	Cys
								130			135		140		
Glu	Thr	His	Leu	His	Trp	His	Thr	Val	Ala	Lys	Glu	Thr	Cys	Ser	Glu
								145			150		155		160
Lys	Ser	Thr	Asn	Leu	His	Asp	Tyr	Gly	Met	Leu	Leu	Pro	Cys	Gly	Ile
								165			170		175		
Asp	Lys	Phe	Arg	Gly	Val	Glu	Phe	Val	Cys	Cys	Pro	Leu	Ala	Glu	Glu
								180			185		190		
Ser	Asp	Asn	Val	Asp	Ser	Ala	Asp	Ala	Glu	Glu	Asp	Asp	Ser	Asp	Val
								195			200		205		
Trp	Trp	Gly	Gly	Ala	Asp	Thr	Asp	Tyr	Ala	Asp	Gly	Ser	Glu	Asp	Lys
								210			215		220		
Val	Val	Glu	Val	Ala	Glu	Glu	Glu	Val	Ala	Glu	Glu				

225 230 235 240
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
 245 250 255
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
 260 265 270
 Ala Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
 275 280 285
 Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
 290 295 300
 Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
 305 310 315 320
 Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
 325 330 335
 Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
 340 345 350
 Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
 355 360 365
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
 370 375 380
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
 385 390 395 400
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
 405 410 415
 Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
 420 425 430
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
 435 440 445
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
 450 455 460
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
 465 470 475 480
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
 485 490 495
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
 500 505 510
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
 515 520 525
 Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
 530 535 540
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
 545 550 555 560
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
 565 570 575
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
 580 585 590
 Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
 595 600 605
 His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
 610 615 620
 Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
 625 630 635 640
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Leu Gly Thr Glu
 645 650 655
 Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu
 660 665 670
 Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro
 675 680 685
 Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys
 690 695 700
 Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys
 705 710 715 720
 Lys Arg Arg Gln Arg Arg Ala His Gln Asn Ser Gln Thr His Gln

725	730	735
Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Ala Asp		
740	745	750
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn		
755	760	765
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn		
770	775	780

<210> 13

<211> 2823

<212> DNA

<213> fusion protein - human

<400> 13

atgctgcccg gttggcaact gtcctgtcg gccgcctgga cggctcgccc gctggaggta 60
cccactgtg gtaatgtcg cctgtggct gaacccaga ttgcattgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtcaagac ccattccccac tttgtgattt cctaccgctg ctttagttgt 360
gagtttataa gtatgtccct tctcggtcct gacaagtgc aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagttag 480
aagagtacca actgtcgatg ctacggcatg ttgtgcctt gccggatgtt caagtccga 540
ggggtagagt ttgtgtgtt cccactggc gaagaaatgtt acaatgttgg ttctgtgtat 600
gcggaggagg atgactcgga tgctgggtt ggcggaggc acacagacta tgcagatggg 660
agtgaagaca aagttagtgcg agtagcagag gaggagaag tggcttgagg ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgtatgagg tagaggaaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtcgttgg aagagggtgt tcgagttcct acaacagcag ccagtacccc ttagtgcgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttgg agggcaagca ccgagagaga atgtcccagg tcatgagaga atggaaagag 1020
gcagaacatgc aagcaaaagaa ctgtccctaa gctgataaga aggcaggatccat ccagcatttc 1080
caggagaaag tggaatctt ggaacaggaa gcggccaaacg agagacagca gctggggag 1140
acacacatgg ccagagtggc agccatgctc aatgaccgccc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc ttttcctcctt cggccctcgatc acgtgttcaa tatgctaaag 1260
aagtatgtcc ggcgagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatgttgg atcccaagaa agccgctcag atccggtccc aggttatgac acacccgt 1380
gtgattttatg agcgtatgaa tcagtcttc tccctgtct acaacgtgcg tgcagtgcc 1440
gaggagattc aggtatgaaatg tgatgagctt cttcggaaatgg acaaaacta ttcatgtac 1500
gtcttggcca acatgatttgc tgaaccaagg atcaggatgc gaaacgtgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccctcccg tgaatggaga gttcagccctg 1620
gacgatctcc agccgtggca ttctttggg gctgactctg tgccagccaa cacagaaaaac 1680
gaagttgagc ctgttgatgc ccgcctgtt gccgaccggag gactgaccac tcgaccagg 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctaga tgcagaattc 1800
cgacatgact caggatataa atgtcatcat caaaaattgg tggttttgc agaagatgtg 1860
ggttcaaca aagggtcaat cattggactc atggggggcg gtgttgtcat agcgcacagt 1920
atcgtcatca ctttgggtat gctgaaagaa aaaaatgtt gtaccggatc cggatccact 1980
agtccagtgt ggtggaaattc tgcatgatatac aagctactgtt cttctatgc acaagcatgc 2040
gatatttggc gactaaaaaa gctcaagtgc tccaaagaaa aaccgaagtgc cgccaaatgt 2100
ctgaagaaca actgggagtg tcgctactt cccaaacca aaaggatctcc gctgactagg 2160
gcacatctga cagaagtggc atcaaggcta gaaagactgg aacagctatt tctactgtt 2220
tttccctcgag aagaccttgc catgatttt aaaaatggatt ctttcatggaa tataaaagca 2280
ttgttaacag gattattttt acaagataat gtatgttggg atgcccgtcac agatagattg 2340
gcttcgttgg agactgtatgc cccctataaca ttggatggc acatgtatgc tgcgacatca 2400
tcatcgaaag agagtatgaa caaaaggatca agacagttga ctgtatcggtt aattccggg 2460
gatctggccc ccccgaccga tgcgttggc gggggacggc tccacttgcgaa cggcgaggac 2520
gtggcgatgg cgcgttggc cgcgttggc gatgttgcgtt tggatgtt gggggacggg 2580
gattccccgg gtcggggatt taccggggatc gactccggcc cctacggcgc tctggatatg 2640
gccgacttgc agtttgagca gatgtttacc gatgccttgc gatgttgc gatcggtggg 2700
gatatccagc acagtggcgg ccggcgttgc gctgttgc accaggagcg ccacctgtcc 2760
aagatcgac agaaccggcata caaaaatccaa acctacaatgtt tctttgagca gatcgac 2820

tag	2823
<210> 14	
<211> 941	
<212> PRT	
<213> fusion protein - human	
<400> 14	
Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala Trp Thr Ala Arg	
1 5 10 15	
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro	
20 25 30	
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln	
35 40 45	
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp	
50 55 60	
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu	
65 70 75 80	
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn	
85 90 95	
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val	
100 105 110	
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu	
115 120 125	
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys	
130 135 140	
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu	
145 150 155 160	
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile	
165 170 175	
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu	
180 185 190	
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val	
195 200 205	
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys	
210 215 220	
Val Val Glu Val Ala Glu Glu Glu Val Ala Glu Val Glu Glu Glu	
225 230 235 240	
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu	
245 250 255	
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile	
260 265 270	
Ala Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg	
275 280 285	
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu	
290 295 300	
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys	
305 310 315 320	
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg	
325 330 335	
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp	
340 345 350	
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu	
355 360 365	
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala	
370 375 380	
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn	
385 390 395 400	
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe	
405 410 415	
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His	
420 425 430	

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
 435 440 445
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
 450 455 460
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
 465 470 475 480
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
 485 490 495
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
 500 505 510
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
 515 520 525
 Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
 530 535 540
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
 545 550 555 560
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
 565 570 575
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
 580 585 590
 Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
 595 600 605
 His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
 610 615 620
 Gly Ala Ile Ile Gly Leu Met Val Gly Val Val Ile Ala Thr Val
 625 630 635 640
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Leu Gly Thr Glu
 645 650 655
 Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
 660 665 670
 Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
 675 680 685
 Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
 690 695 700
 Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
 705 710 715 720
 Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
 725 730 735
 Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
 740 745 750
 Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
 755 760 765
 Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
 770 775 780
 Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
 785 790 795 800
 Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
 805 810 815
 Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp
 820 825 830
 Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala
 835 840 845
 Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly
 850 855 860
 Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met
 865 870 875 880
 Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp
 885 890 895
 Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Ala Asp Ala Ala
 900 905 910
 Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr
 915 920 925

Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
930 935 940

<210> 15
<211> 2823
<212> DNA
<213> fusion protein - human

<210> 16
<211> 941
<212> PRT

<213> fusion protein - human

<400> 16

Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1 5 10 15
 Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
 20 25 30
 Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
 35 40 45
 Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
 50 55 60
 Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
 65 70 75 80
 Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
 85 90 95
 Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
 100 105 110
 Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
 115 120 125
 Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
 130 135 140
 Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
 145 150 155 160
 Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
 165 170 175
 Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
 180 185 190
 Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
 195 200 205
 Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
 210 215 220
 Val Val Glu Val Ala Glu Glu Glu Val Ala Glu Val Glu Glu Glu
 225 230 235 240
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
 245 250 255
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
 260 265 270
 Ala Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
 275 280 285
 Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
 290 295 300
 Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
 305 310 315 320
 Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
 325 330 335
 Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
 340 345 350
 Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
 355 360 365
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
 370 375 380
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
 385 390 395 400
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
 405 410 415
 Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
 420 425 430
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
 435 440 445
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
 450 455 460
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala

465	470	475	480
Glu	Glu	Ile	Gln
Gln	Asp	Glu	Val
Glu	Asp	Leu	Gln
485	490	495	
Tyr	Ser	Asp	Asp
Asp	Val	Leu	Ala
Asn	Met	Ile	Ser
500	505	510	
Tyr	Gly	Asn	Asp
Asp	Ala	Leu	Met
Pro	Ser	Leu	Thr
515	520	525	
Val	Glu	Leu	Leu
Pro	Leu	Pro	Val
Asn	Gly	Glu	Phe
530	535	540	
Pro	Trp	His	Ser
Phe	Gly	Ala	Asp
Ser	Val	Pro	Ala
545	550	555	560
Glu	Val	Glu	Pro
Pro	Val	Asp	Ala
Asp	Arg	Pro	Ala
565	570	575	
Thr	Arg	Pro	Gly
Ser	Gly	Leu	Thr
580	585	590	
Glu	Val	Lys	Met
Asp	Ala	Glu	Phe
595	600	605	
His	His	Gln	Lys
610	615	620	
Gly	Ala	Ile	Ile
Gly	Leu	Met	Val
Gly	625	630	635
Ile	Val	Ile	Thr
Leu	Val	Met	Leu
Lys	640	645	650
Ile	Val	Ile	Thr
Leu	Val	Met	Leu
Lys	655	660	665
Leu	Gly	Ser	Thr
Ser	Pro	Val	Trp
660	665	670	
Leu	Ser	Ser	Ile
Ile	Glu	Gln	Ala
Cys	Asp	Ile	Cys
675	680	685	
Lys	Cys	Ser	Lys
Glu	Lys	Pro	Lys
Cys	Ala	Cys	Lys
690	695	700	
Trp	Glu	Cys	Arg
Arg	Tyr	Ser	Pro
705	710	715	720
Ala	His	Leu	Thr
Glu	Val	Glu	Ser
725	730	735	
Phe	Leu	Leu	Ile
Phe	740	745	750
Asp	Ser	Leu	Gln
755	760	765	
Asp	Asn	Val	Asn
770	775	780	
Thr	Asp	Met	Pro
785	790	795	800
Ser	Ser	Glu	Gly
Ser	Ser	Ser	Asn
805	810	815	
Gly	Ile	Pro	Gly
Asp	Leu	Ala	Pro
820	825	830	
Glu	Leu	His	Leu
Asp	Gly	Glu	Asp
835	840	845	
Leu	Asp	Asp	Phe
850	855	860	
Pro	Gly	Phe	Thr
865	870	875	880
Ala	Asp	Phe	Glu
885	890	895	
Glu	Tyr	Gly	Asp
900	905	910	
Val	Thr	Pro	Glu
915	920	925	
Glu	Asn	Pro	Thr
930	935	940	

<211> 3354
<212> DNA
<213> human

<400> 17

agtttcctcg gcagcggtag gcgagagcac gcccgggggc gtgcgcgggg gccccgggag 60
acggccggcg tggcgccgca ggcagagcaa ggacgcggcg gatcccactc gcacagcagc 120
gcactcggt ccccgccag ggtcgcatg ctgcccgggtt tggactgct cctgctggcc 180
gcctggacgg ctggggcgct ggaggtaccc actgatggta atgctggccct gctggctgaa 240
ccccagattt ccatgttctg tggcagactg aacatgcaca tgaatgtcca gaatggaaag 300
tgggatttcg atccatcagg gaccaaaacc tgcatgtata ccaaggaaagg catcctgcag 360
tattccaaag aagtctcagg tgaactgcag atccacaat tggtagaaagg caaccaacca 420
gtgaccatcc aagaactgggtt caagggggcc cgcaaggatg gcaagagccca ccccccattt 480
gtgattccct accgcgtctt agttggtag tttgttaatgt atgccttctt cgttctgac 540
aagtgc当地 tcttacacca ggagaggatg gatgttgcg aaactcatct tcactggcac 600
accgtcgcca aagagacatg cagtgagaag agtaccaact tgcatgacta cggcatgttg 660
ctgcccctcg gaattgacaa gttccggaggg gtagagttt tggtagttcccc actggctgaa 720
gaaagtgaca atgtggattt tgctgtatgcg gaggaggatg actcggatgt ctgggtgggc 780
ggagcagaca cagactatgc agatgggagt gaagacaaag tagtagaaatg agcagaggag 840
gaagaagtgg ctgaggttga agaagaagaa gccgtatgt acgaggacca tgaggatgtt 900
gatgaggtttaggaaaccc tacgaagaag ccacagagaa aaccaccaggc 960
attgccacca ccaccacccac caccacagag tctgttggaaatg aggtggatcg agttcttaca 1020
acagcagcca gtaccctgtg tgccgttgc aagtatctcg agacacctgg gnatgagaat 1080
gaacatgccc atttccagaa agccaaagag aggcttgagg ccaagcaccg agagagaatg 1140
tcccaggctca tgagagaatg ggaagaggca gaacgtcaag caaaagaactt gcctaaagct 1200
gataagaagg cagtatccca gcatttccag gagaatgtgg aatctttgga acaggaagca 1260
gccaacgaga gacacgactt ggtggagaca cacatggca gagtggaaatg catgtcaat 1320
gaccggccgc ccctggccctt gggaaatc acatccgtctc tgccgttgcg 1380
cctcgtcactg ttgttcaatat gttttttttt gatgttgcgc atggggatc ccaagaaaatg 1440
cacaccctaa agcatttgcg gcatgttgcgc ttatattttttt gatgttgcgcg 1500
cggttccagg ttatgacaca cctccgttgc atttatgatc gcatgaatca gttcttcc 1560
ctgctctaca acgttgcctgc agtggcccgag gagattcagg atgaagtgtt tgagctgtt 1620
cagaaaagac aaaactattt agatgtacgtc ttggccaaca tgattttttt gatgttgcgc 1680
agttaacggaa acgtatgtct catgtccatct ttggccaaaaa cggaaaaccac cgtggagctt 1740
cttccctgttga atggagatgtt cggccgttgc gatcttccgc cgtggccattt tttttttttt 1800
gactctgtgc cggccacac agaaaacggaa ttgtggatc ttgttgcgcg 1860
gaccggggac tgaccactcg accaggatctt gggggatc atatcaagac ggaggagatc 1920
tctgaagtta agatggatgc agaattccga catgacttagt gatatgttgc gatatgttgc 1980
aaattttttt tcttttgcaga agatgtgggt tcaaaacaaatg gtgcaatcat tggactcatg 2040
gtggccgggtt ttgttcatatc gacagtgtatc gtcatcacct ttgttgcgcg 2100
cagtacacat ccatttcatca ttgttgcgcg ggggttgcgcg 2160
cgccacccgtt ccaatgtca gcaacccggc tacggaaatc caacccatcaatc gttcttgc 2220
catgttgcaga actagacccca ccccaacaccc gcttcttgcg ttggacaccaaaatc 2280
ttcaatcccc atcgttgcgcg attttttttt taatgttgcgcg 2340
atttacttcat tatgttgcgcg ttacagatgtt gctgttgcgcg 2400
attaatccac acatcatgttgc ttgttgcgcg 2460
ttattaaatgg gtttttgcgcg ttgttgcgcg 2520
atttcttctt gattttttt catgttgcgcg 2580
tgtgacccaa ttaatgttgcgcg ttgttgcgcg 2640
tgaacgttgcg agtttgcgcg ttgttgcgcg 2700
ttttttttttaa atacccatgttgcgcg 2760
gcattttact gtacatgttgc ttgttgcgcg 2820
cacacgttgcg ttgttgcgcg ttgttgcgcg 2880
ttttttttttaa atacccatgttgcgcg 2940
gttggacttgc ttacggggcg gttggggagg ggttgcgcg 3000
ttctccaaaa caatgttgcgcg caggatgttgc ttgttgcgcg 3060
tttcttactat gtattacata aataaattaa ataaaataatc cccggggcaag 3120
gaaggatgac tacagacattt aataaattatc agtaatttt gttggggagg 3180
tcaatttttt ttaacccatgttgcgcg ttgttgcgcg 3240
gttggacttgc ttacggggcg gtaaggatgttgcgcg 3300
ttcaattttttaa ataaaatgttgcgcg aaataaataatc attttttttt aatgttgcgcg 3354

<210> 18
<211> 695
<212> PRT
<213> human

<400> 18
Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
100 105 110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
115 120 125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
130 135 140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145 150 155 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
165 170 175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
180 185 190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
195 200 205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
210 215 220
Val Val Glu Val Ala Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225 230 235 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
245 250 255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
260 265 270
Ala Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Val Val Val Arg
275 280 285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
290 295 300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305 310 315 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
325 330 335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
340 345 350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
355 360 365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
370 375 380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385 390 395 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
405 410 415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
420 425 430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
435 440 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
 450 455 460
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
 465 470 475 480
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
 485 490 495
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
 500 505 510
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
 515 520 525
 Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
 530 535 540
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
 545 550 555 560
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
 565 570 575
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
 580 585 590
 Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
 595 600 605
 His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
 610 615 620
 Gly Ala Ile Ile Gly Leu Met Val Gly Val Val Ile Ala Thr Val
 625 630 635 640
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
 645 650 655
 His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
 660 665 670
 His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
 675 680 685
 Phe Phe Glu Gln Met Gln Asn
 690 695

<210> 19
 <211> 3972
 <212> DNA
 <213> human

<400> 19
 atggtagaca agggcgagga gctgttcacc ggggtggtgc ccatacctgggt cgagctggac 60
 ggcgactaa acggccacaa gttcagctgt tccggcgagg gcgaggcgca tgccacctac 120
 ggcaagctga cccctgaagtt catctgcacc accggcaagc tgccccgtgcc ctggcccacc 180
 ctcgtgacca ccttcaccta cggcgtgcag tgcttcgtccc gctaccccgaa ccacatgaag 240
 cagcacgact tcttcaagtc cggccatgccc gaaggctacg tccaggagcg caccatcttc 300
 ttcaaggacg acggcaacta caagaccgcg gccgagggtga agttcgaggcg cgacaccctg 360
 gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420
 aagctggagt acaactacaa cagccacaag gtctatatca ccggccgacaa gcagaagaac 480
 ggcatcaagg tgaacttcaa gacccgcac aacatcgagg acggcagcg tgctcgcc 540
 gaccactacc agcagaacac ccccatcgcc gacggccccc tgctgctgcc cgacaaccac 600
 tacctgagca cccagtccgc cctgagcaaaa gaccccaacg agaaggcgca tcacatggc 660
 ctgctggagt tcgtgaccgc cgccgggatc actctcgca tggacgagct gtacaagtaa 720
 ctcgagtcta gagggcccg ttaaaccgcg tgatcagcct cgactgtgcc ttctagttgc 780
 cagccatctg ttgtttgccc ctcccccggt ctttccttga ccctggaaagg tgccactccc 840
 actgtccttt cctaataaaaa tgaggaaatt gcatcgatt gtctgagtag gtgtcattct 900
 attctggggg gtgggggtggg gcaggacacg aaggggggagg attgggaaga caatagcagg 960
 catgctgggg atgcgggtggg ctctatggct tctgaggcgaa aagaaccag catgtgagca 1020
 aaaggccacg aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg 1080
 ctccgcggcc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaaccac 1140
 acaggactat aaagatacca ggcgttcccc cctgaaagct ccctcgatcg ctctcctgtt 1200
 ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcggaaag cgtggcgctt 1260
 tctcatacgct cacgctgttagt gtatctcagt tcgggttagg tcggtcgctc caagctggc 1320

tgtgtgcacg aaccggccgt tcagccccac cgctgcgcct tatccggtaa ctatcgctt 1380
 gagtcacaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt 1440
 agcagagcga ggtatgttagg cggtgctaca gagttcttga agtggggc taactacggc 1500
 tacactagaa gaacagtatt tggtatctgc gctctgctga agccaggatc cttcgaaaaa 1560
 agagttggta gctcttgatc cggcaaacaa accaccgctg gttagcggtgg ttttttgtt 1620
 tgcaagcagc agattacgcg cagaaaaaaaaa ggatctcaag aagatcctt gatctttct 1680
 acggggctcg acgcctagtg gaacgaaaaac tcacgtaag ggattttgtt catgacattt 1740
 acctataaaaa ataggcgtat cacgaggccc tttcgctcg cgcgtttcg tgatgacggt 1800
 gaaaacctct gacacatgca gctcccgag acggcacag cttgtctgtt agcggatgcc 1860
 gggagcagac aagcccgta gggcgcgta gcggtgtt gcggtgtt gggctggctt 1920
 aactatcgcc catcagagca gattgtactg agatgcacc atatcggtt tgaaataccg 1980
 cacagatcgcc taaggagaaaa ataccgcac aggacgcgc ctgtagcggc gcattaagcg 2040
 cggcgggtgt ggtgttacg cgcagcgtga ccgctacact tgccagcgc ctagcgcgg 2100
 ctcccttcgc ttcttcctt tccttctcg ccacgttcgc cggctttccc cgtcaagctc 2160
 taaatcgcccc gctccctta gggttccgat ttatgtctt acggcacctc gacccaaaaa 2220
 aacttgatta ggggtatggt tcacgtatg ggcacatcgcc ctgatagacg gtttttcgccc 2280
 ctttgacgtt ggagtccacg ttctttaata gtggacttctt gttccaaact ggaacaacac 2340
 tcaaccctat ctccgttatc tcttttattataaaggat ttgcccattt tcggcttatt 2400
 gttaaaaaaaaa tgatctgtt taaacatgca ttaacgcgaa ttttacaaaaa attataaacgc 2460
 ttacaatttc cattcgccat tcaggctgaa cttagatctag aatccgttac ataacttacg 2520
 gtaaatggcc cgccctggctg tatgttccca tagtaacgcg accggccaaac gaccccccgc 2580
 cggtaaaactg cccacttggc aatagggact ttccattgac gtcaatgggt ggagtattt 2640
 gacgtcaatg acggtaaatg agtacatcaa gtgtatcata tgccaagtac gccccctatt 2700
 tttcttactt ggcagtacat ctacgtatcgtt gtcacgtcta tattatggc agtacatgac 2760
 tggcgttaca tcaatggcg cccattgacg tcaatgggag cgtaacaact cggccccattt 2820
 gtcgccccgg gggctggatc tggatagccg tttgactcac ggggatttcc aagtctccac 2880
 gacgcaaatg ggcggtaggc gtgtacgggt ggaggcttat 2940
 gacccgtcag atccgttgcgaa gacccatcc acgctttt 3000
 gacccgtcag atccgttgcgaa gacccatcc acgctttt 3060
 ggaccgatcc agccctcccgcc gccggaaacg gtgcatttgcgaa 3120
 atcatcgccat tagtataatcg gcatacgatata atacgacaag 3180
 caagttgacc agtgcgcgtt cgggtgcctac cgcgcgcgc 3240
 ctggaccgac cggctcggtt tctccgggggatc cttcggtgg 3300
 gacgacttcg cgggtgttgcgac gtgaccctgt tcacgcgcg 3360
 cctggcctgg gttgggtgc gccggcttgcgac gtagctgtac 3420
 gtccacgac ttccggacg cctccggggcc ggcacatgacc 3480
 gggccgggag ttccggctgc ggcacccggc cggcaactgc 3540
 gtgcacttcg tggccgagga gcggactgtc cactcgaccc ttttgcgc 3600
 ttataatgtt tacaataaaaa gcaatagcat cacaatattc acaaaaaaaatc catttttttc 3660
 actgcatttctt agttgtggtt tgcggacttactt catcaatgtt tctggatcg 3720
 tacttcaaga actgtcgaca tcgacttgc tacaatggggatc ttcggctgg 3780
 gggaggcgtg gcctggggcg gactggggatc tggcggccccc ttagatccttgcatataagca 3840
 gctgtttttt gcctgtactg ggtctctcg gtttagaccatctgagctctc 3900
 tggctacta gggaaacccac tgcttaagcc tcaataaaagc ttggtagccgac gctcgatcc 3960
 gaattcggcca cc 3972

<210> 20
 <211> 3972
 <212> DNA
 <213> human

<400> 20
 taccactctgt tcccgctctt cgacaagtgg cccacccacg ggtaggacca gtcgacccgt 60
 cccgtcgatc tggccgtgtt caatgtcgac aggccgcgtcc cgctcccgatc acgggtggatc 120
 ccgttcgact gggacttcaa gtagacgtgg tggccgttgc acggggcacgg gaccgggtgg 180
 gagcacttggt ggaagtggat cccgcacgtc acgaaggccgg cgatggggct ggtgtacttc 240
 gtcgtctgtt agaaatgttgc gccgtacggg cttccgtatgc aggtccctgc gtggtagaaag 300
 aagttcttcg tggccgttgc gttctggggc cggctccact tcaagctccc gctgtgggac 360
 cacttggcgt agctcgactt cccgtacgtt aagttctcc tggccgttgc ggcacccgtg 420
 ttgcacccatc tgggtatgttgc gtcgggttgc cagatataatgc ggcggctgtt cgtcttcttgc 480
 ccgttagtcc acttgaagtt ctggccgttgc ttgttagctcc tggccgtcgcgac gtcgagcgg 540
 ctggtagtgg tggccgttgc gggtagccg ctggccggggc acgacgacgg gctgtgggtg 600

atggactcgt gggtcaggcg ggactcggtt ctggggttgc tcttcgcgt agtgtaccag 660
 gacgacctca agcaactggcg gcggccctag tgagagccgt acctgctcg catgttcatt 720
 gagctcagat ctccccggca aatttggcg actagtccga gctgacacgg aagatcaacg 780
 gtcggtagac aacaaacggg gaggggcac ggaaggaaact gggaccttc acggtgaggg 840
 tgacagaaaa ggattatttt actcctttaa cgtacgcgtaa cagactcatc cacagtaaga 900
 taagaccccc cacccccaccc cgtcctgtcg ttccccctcc taacccttct gttatcgtcc 960
 gtacgacccc tacgcccaccc gagataccga agactccgcc tttcttggc gtacactcg 1020
 tttccgtcgt tttccggc cttggcattt ttccggcga acgaccgc aaaggatattc 1080
 gaggcggggg gactgctcg agtgtttta gctggcagggt cagtctccac cgcttggc 1140
 tgtccgtata tttctatgtt ccgcaaaaggg ggacccctcg gggagcacgc gagaggacaa 1200
 ggctggacg gcgaatggcc tatggacagg cgaaagagg gaagggccctc gcaccggc 1260
 agagtatcga gtgcgacatc catagagtca agccacatc agcaagcgat gttcgacccg 1320
 acacacgtgc ttggggggca agtcgggctg ggcacccggaa ataggccatt gataggcagaa 1380
 ctcagggtgg gcattctgt gctgaatagc ggtgaccgtc gtcggtgacc attgtcctaa 1440
 tcgtctcgct ccatacatcc gccacgtgt ctcaagaact tcaccaccgg attgtatgcc 1500
 atgtatctt cttgtcataa accatagac cgagacgact tcggtaatg gaagccttt 1560
 tctcaaccat cgagaactag gccgtttttt tggtggcgc acatggccacc aaaaaaaaaa 1620
 acgttcgtcg tctaattgcg cttttttttt cctagagttc ttcttagaaaa cttagaaaa 1680
 tgccccagac tgcgactac ctgtttttt agtgcatttcc cttaaaacca gtaactgtaat 1740
 tggatatttt tattccgcata gtgctccggg aaaggcagac ggcggaaaggcc actactgcca 1800
 cttttggaga ctgtgtacgt cgagggcctc tgccagtgtc gaacagacat tcgcctacgg 1860
 ccctcgtctg ttccggcagt cccgcgcgt cgcccaacaac cgcccacagc cccgaccgg 1920
 ttgatacggcc gtagtctcg ttaacatgac tctcacgtgg tatacggccac actttatggc 1980
 gtgtctacgc attccctttt tatggcgtag tccttcgcgg gacatcgccg cgtaattcgc 2040
 gcccggccaca ccaccaatgc gcgtcgcact ggcgtatgtg acgggtcgccg gatcgccggc 2100
 gaggaaaggcg aaaggaaaggc aggtcaaggc gccgaaagggg gcagttcgag 2160
 atttagcccc cggggaaatc cccaagggtt aatcacgaaa tgccgtggag ctgggtttt 2220
 ttgaactaat cccactacca agtgcatac cccgttagccg gactatctgc caaaaaaaa 2280
 gaaactgcaa ccttcagggtc aagaaaattt cacctgagaa caagggttga cttttttgt 2340
 agttgggata gaggcagata agaaaactaa atattcccta aaacggctaa agccggataa 2400
 ccaattttt actcgactaa attgtttttt aattgcgtt aaaattgttt tataattgcg 2460
 aatgttaaag gtaagcggtt agtccgactt gatcttagatc tcaggcaatg tattgaatgc 2520
 catttacccgg gcgaccggac tggcgggttgc tggggccgg gtaactcgat ttattactgc 2580
 atacaagggt atcattcggg ttatccctga aaggttaactg cagttacccca cctcataaat 2640
 gccatttgac gggtaaccgc tcatgttagtt cacatagtat acggttcatg cgggggataa 2700
 ctgcaggattac tgccatttac cggccggacc gtaatacggg tcatgtactg gaataccctg 2760
 aaaggatgaa ccgtcatgtat gatgcataat cagtagcgat aatggtacca ctacccaaa 2820
 accgtcatgt agttacccgc acctatcgcc aaacttagtgc cccctaaagg ttcaagggtg 2880
 gggtaactgc agttacccctc aaacaaaacc gtgggttttag ttgcctctgaa aggttttaca 2940
 gcatgttga ggggggttaa ctgcgttac ccgcacatccg cacatggccac cctccagata 3000
 tattcgctctc gggccatataa ctggcgcgt tagggacact ctgcggtagg tgccgacaaaa 3060
 ctggagggtat ctctgtggc cctggcttagg tcggaggcgc cggcccttgc cacgtaacct 3120
 tgccctggcac aactgttaat tagtagccgt atcatatagc cgtatcatat tatgtctgtc 3180
 cactccgttga ttgggtaccg gttcaactgg tcacggcaag gccacgagtg gcgcgcgcgt 3240
 cagccggctc gccagctcaa gacctggctg gccgagccca agagggccct gaagcacctc 3300
 ctgctgaagc gggccacacca gggccctgtc cactgggaca agtagtgcgc ccaggccctg 3360
 gtccaccacg gcctgtgtg ggaccggacc cacacccacg cggccggacatg gtcgacatg 3420
 cgggtccacca gcctccaggca cagggtcttgc aaggccctgc ggaggccccc cgggtactgg 3480
 ctcttagccgc tcgtcggcac cccggccctc aagccggacgc cgctggccgc gccgttgacg 3540
 cacgtgaagc accggctctc cgtcctgact gtgagctggg gctttgaaca aataacgtcg 3600
 aatattacca atgtttttt cgttatacgta gtgtttaaag tggttatttc gtaaaaaaa 3660
 tgacgttaaga tcaacaccaa acagggttgc gtagttacat agaatagtac agacctagct 3720
 atgaagttct tgacgtactgt agctcgaaacg atgttccctg aaaggcgacc cctgaaagg 3780
 ccctccgcac cggaccggcc ctgacccttc accgctccgg agtcttaggac gtatattcg 3840
 cgacgaaaaa cggacatgac ccagagagac caatctggc tagactcgga ccctcgagag 3900
 accgattgt cccttgggtg acgaattcgg agttatttcg aaccatggct cgagcttagg 3960
 cttaaaggcg gg 3972

<210> 21
 <211> 7153
 <212> DNA
 <213> human

<400> 21
gacggatccg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aaggcagtat ctgctccctg cttgtgtt ggaggtcgct gtagtagtcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgcaccga caattgcatt aagaatctgc 180
ttagggttag cgctttgcg ctgcttcgcg atgtacgggc cagatatacg cggtgacatt 240
gattttagc tagtattaa tagtaatcaa ttacgggtc attagttcat agcccatata 300
tggagttccg ctgttacataa ttacggtaa atggccgcg tggctgaccg cccacgacc 360
cccgccttatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggacttcc 420
attgacgtca atgggtggac tatttacggt aaactgccc cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaaatggccc gcctggcatt 540
atgcccagta catgaccta tgggacttc ctacttggca gtacatctac gtattagtca 600
tcgcttattc catggtgatg cggttttggc agtacatcaa tgggcgttga tagcgggtt 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tggaggttgg tttggcacc 720
aaaatcaacg ggacccatca aaatgtcgt aacaactccgc cccattgacg caaatggcg 780
gtagggcgtgt acgggtggag gtcttatataa gcagagctct ctggcttaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggcttagc 900
gtttaaactt aagcttcccc gcgcagggtc gcgatgctgc cccggttggc actgctctg 960
ctggccgcct ggacggctcg ggcgctggag gtacccactg atggtaatgc tggcctgctg 1020
gctgaacccc agattgccc gttctgtggc agactgaaca tgcacatgaa tgcacatgaa 1080
gggaagtggg attcagatcc atcagggacc aaaacctgca ttgataccaa ggaaggcattc 1140
ctgcagtatt gccaagaagt ctaccctgaa ctgcagatca ccaatgttgc agaaggcacc 1200
caaccaggta ccatccagaa ctggtgcaag cggggccgca agcagtgcgaa gaccatccc 1260
cactttgtga ttcccttaccg ctgctttagt ggtgagttta taagtgtatgc ctttctcg 1320
cctgacaagt gcaatttctt acaccaggag aggatggatg tttgcgaaac tcatcttcac 1380
tggcacaccc tcgccaaaga gacatgcagt gagaagagta ccaacttgcg tgcactacggc 1440
atgttgcgc cctgcggaaat tgacaagttc cgaggggttag agtttgttgc ttgcccactg 1500
gctgaagaaa gtgacaatgt ggattctgt gatgcggagg aggatgactc ggtatgtctgg 1560
tggggcggag cagacacaga ctatgcagat gggagtgaaagccaaatgttgcgatgca 1620
gaggaggaag aagtggctga gttggaaagaa gaagaagccg atgatgcgaa ggacgttag 1680
gatgtgtatc agtagagggatc agggcttag gAACCTACAG aagaagccac agagagaacc 1740
accaggattt ccaccacccac caccacccacc acagagtctg tggaaagaggt gttcgagtt 1800
cctacaacacg cagccagtac ccctgtatgcc gttgacaagt atctcgagac acctgggat 1860
gagaatgaac atgcccattt ccagaaaaggc aaagagaggc ttgaggccaa gcaccgagag 1920
agaatgtccc aggtcatgag agaatggaa gaggcagaac gtcaagcaaa gaacttgcct 1980
aaagctgata agaaggcagt tatccagcat ttccaggaga aagtggaaatc ttggaaacag 2040
gaagcagcca acggagagaca gcagctggg gagacacaca tggccagat ggaaggcattc 2100
ctcaatgacc gcccgcgc ggccttggg aactacatca ccgtctcgca ggctgttct 2160
cctcgccctc gtacgtgtt caatatgcta aagaagtatg tcccgccaga acagaaggac 2220
agacagcaca ccctaaagca tttcgagcat gtgcgcattt tgatcccaaa gaaagccct 2280
cagatccgtt cccaggttat gacacacccctc cgtgttattt atgagcgcattt gaatcgtct 2340
ctctccctgc tctacaacgt gcttgcagtg gccgaggaga ttcaaggatgagttt gttgtatg 2400
ctgttccaga aagggaaaaat ctattcagat gacgttctgg ccacatgtat tagtgaacca 2460
aggatcgtt acggggaaacgta tgcttcctat ccattttgtt ccggaaacgcgaa aaccacgtt 2520
gagctccctc ccgttgcattt agagttcgcg ctggcgcattt tccaggccgtg gcattttttt 2580
ggggctgtact ctgttgcgcg caacacagaa aacgaagttt agcgttgcgatg tggccgcctt 2640
gcttgcgcacc gaggactgac cactcgacca gtttctgggt tgacaaatataa caagacggag 2700
gagatcttg aagtgaatct agatgcagaa ttccgcacatg actcaggata tgaagttcat 2760
cataaaaat tgggtttctt tgcagaagat gtgggttcaa acaaagggtc aatcatggaa 2820
ctcatgttgg gccgttgcgtt catagcgaca gtatgcgtca tcaaccttgcgatgatgttgc 2880
aagaagata tcatggggcc agtagatccct agactagacg ccttggaaagca tccaggaaatg 2940
cagcttaaaa ctgttgcattt caatttgcattt tgtaaaaatgttgcgttca ttggcaattt 3000
tggttcatga caaaaggctt aggcatctcc tatggcaggaa agaaggcggag acagcgcacg 3060
agagctcatc agaacagtca gactcatcaa gcttctctat caaagcagta agtagggcgc 3120
cgctcgagtc tagagggccc gtttaaaccctt gctgatcaggc ctgcactgttgc ctttctgtt 3180
gccagccatc tggttgc ecctcccccc tgccttcctt gaccctggaa ggtgccactc 3240
ccactgttctt tccctaataa aatgaggaaa ttgcacatgc tttgttgcgtt aggtgttcat 3300
ctattctggg ggggggggtt gggcaggacaa gcaagggggaa ggattgggaa gacaatagca 3360
ggcatgttgg ggtatggcgtt ggcttctatgg ctgttgcattt ggaaagaacc agtggggctt 3420
ctaggggtt tccccacccgc ccctgttagcg gcgcattaaag cggccgggggtt gtgttgcgtt 3480
cgccgcgggtt gaccgttca cttggccaggc ccctagcgc cgcgttcccttgc tttttttttt 3540
cttccttttcc cgccacgttc gccggctttt cccgtcaaggc tctaaatcggtt ggcacccctt 3600

tagggttccg atttagtgc ttacggcacc tcgaccccaa aaaacttgat tagggtgatg 3660
 gttcacgtg tggccatcg ccctgataga cgggtttcg cccttgacg ttggagtcca 3720
 cgttcttaa tagtgactc ttgttccaaa ctggaaacaac actcaaccct atctcggtct 3780
 attctttga tttataaggg attttggga tttccgccta ttggttaaaa aatgagctga 3840
 tttaaacaaa atttaacgcg aattaattct gtggaatgtg tgcagttag ggtgtggaaa 3900
 gtccccaggc tccccaggca ggcagaagta tgcaagcat gcatctcaat tagtcagcaa 3960
 ccagggtgtgg aaagtccccca ggctccccag cagggcagaag tatgcaaagc atgcatctca 4020
 attagtgcg aaccatagtc ccgcccccaa ctcccccata cccgcccccta actccgccc 4080
 gttccgccc ttctccgccc catggctgac taattttttt tatttatgca gaggccgagg 4140
 ccgcctctgc ctctgagcta ttccagaagt agtggaggagg ctttttggg ggcctaggct 4200
 ttgcaaaaaa gctccggga gcttgtatata ccattttcg actgtatcag cacgtgtga 4260
 caattaatca tcgcatata atatcgccat agtataatc gacaagggtga ggaactaaac 4320
 catggccaag ttgaccagtg ccgttccggt gctcacccgcg cgcgcacgtc ccggagcgg 4380
 cgagttctgg accgaccggc tcgggttctc ccggacttc gtggaggacg acttcgccc 4440
 tgtggtccgg gacgacgtga ccctgttcat cagcgcggc caggaccagg tgggccgga 4500
 caacaccctg gcctgggtgt ggggtgcgccc cctggacgag ctgtacgccc agtggtcgga 4560
 ggtcggtc acgaacttcc gggacgcctc cggggccggc atgaccgaga tcggcgagca 4620
 gccgtgggg cggaggttcg ccctgcgcg cccggccggc aactgcgtc acttcgtggc 4680
 cgaggagcag gactgacacg tgcgtacgaga ttgcgttcc accgcgcctc tctatgaaag 4740
 gttgggttc ggaatcggtt tccggacgc cggctggatg atcctccagc gcggggatct 4800
 catgctggag ttctccgccc accccaaactt gtttattgca gttataatg gttacaaata 4860
 aagcaatagc atcacaattt tcacaaataa agcattttt tcaactgcatt ctatgttgg 4920
 tttgtccaaa ctcatcaatg tatcttatca tgcgtgtata cgcgtaccc ttagcttagag 4980
 ctggcgtaa tcatggtcat agctgttcc tgcgtgaaat ttttatccgc tcacaattcc 5040
 acacaacata cgagccggaa gcataaaatg taaagcctgg ggtgcctaat gagtgagcta 5100
 actcacatta attcgtgtc gtcactgccc cgcttccag tggggaaacc tgcgtgc 5160
 gctgcattaa tgaatcgccca aacgcgcggg gagaggcggt ttgcgtattt ggcgccttc 5220
 cgcttcctcg ctcaactgact cgctgcgtc ggtcggtcg ctgcggcggag cggatcagc 5280
 tcactcaaag gcggtaaatac ggttatccac agaatcaggg gataacgcag gaaagaacat 5340
 gtgagcaaaa ggccagcaaa agggcagggaa ccgtaaaaaaag gccgcgttgc tggcggtttt 5400
 ccatagcgtc cgccccccctg acgagcatca caaaaatcga cgctcaagtc agagggtggcg 5460
 aaaccggaca ggactataaa gataccggc gttttccctt ggaactcc tcgtgcgtc 5520
 tcctgttccg acctcgtccg ttacccggata cctgtccgc tttctccctt cgggaagcgt 5580
 ggcgtttct caatgctcac gtcgttagta ttcgtatcgtc gtcgttagtgc ttgcctccaa 5640
 gctgggtctgt gtgcacgaac ccccggttca gcccggaccgc tgcgccttat ccggtaacta 5700
 tcgtcttgag tccaaacccgg taagacacga cttatgcaca ctggcagcag ccactggtaa 5760
 caggattagc agagcgaggt atgtaggcgg tgcgtacagag ttcttgaagt ggtggctaa 5820
 ctacggctac actagaagga cagtattttg tatctgcgtc ctgcgtgaagc cagttaccc 5880
 cggaaaaaaa gttgttagct cttgtatccg caaacaaacc accgcgttgc gcggtggttt 5940
 ttttgttgc aagcagcaga ttacgcgcag aaaaaggaa ttcgtatcgtc gtcgttagtgc 6000
 ctttctacg gggctgtacg ctcgtggaa cggaaactca cgttaagggtt ttttgtcat 6060
 gagattatca aaaaaggatct tcaccttagat ccttttaaat taaaatgaa gttttaaatc 6120
 aatctaaatg atatatgagt aaacttggtc tgacagttaa caatgtttaa tcagtggagg 6180
 acctatctca gcgatctgtc ttttcgttc atccatagtt gcctgactcc ccgtcggtta 6240
 gataactacg atacgggagg gcttaccatc tggcccccagt gtcgtatcgtc taccgcgaga 6300
 cccacgctca ccgcgtcccg atttacatc aataaaccag ccagccggaa gggccgagcg 6360
 cagaagtggt cctgtcaactt tatccgcctt catcgtatcgtc attaattttt gccggaaaggc 6420
 tagagtaatg agttcgccag ttaatagttt ggcgtatcgtc gtcgttagtgc ttcacaggcat 6480
 cgtgggtctca cgctcggtcg ttggatgtgc ttccatcgtc tccgggttccca aacgtatcgtc 6540
 gcgagttaca tgatccccca tgcgtatcgtc aaaagggtt agtcgtatcgtc gtcgtatcgtc 6600
 cgtgtcaga agtaagttgg ccgcgtatcgtc atcactcatg gttatggcag cactgcataa 6660
 ttctcttact gtcatggcat ccgtatcgtc cttttttttt actgggtgatg actcaaccaa 6720
 gtcattctga gaatagtgtt tgccggacc gagggtgtct tgcgtatcgtc gtcgtatcgtc 6780
 taataccgcg ccacatcgca gaaactttaaa agtgcgtatcgtc ttcacaggcat gtcgtatcgtc 6840
 gcgaaaactc tcaaggatct taccgtatcgtc gagatccgt tgcgtatcgtc ccactcggtc 6900
 acccaactga tcttcgtatcgtc cttttttttt caccatcggtt tgcgtatcgtc gtcgtatcgtc 6960
 aaggcaaaaat gcccggaaaaa aggaaataag ggcgtatcgtc aatgttggaa tactcataact 7020
 ctccctttt caatattttt gtcgtatcgtc tgcgtatcgtc gtcgtatcgtc gtcgtatcgtc 7080
 atttgaatgt atttagaaaaa ataaaacaaat aggggttccg cgcacatttc cccgaaaaagt 7140
 gccacactgac gtc 7153

<210> 22
<211> 7153
<212> DNA
<213> human

<400> 22

ctgcctagcc ctctagaggg ctaggggata ccagctgaga gtcatgttag acgagactac 60
ggcgatataaa ttccggtcata gacgaggggac gaacacacaa cctccagcga ctcacacgc 120
gctcgatgtt aattcgatgt tttccgttc cgaactggct gtaacgtac ttcttagacg 180
aatccccatc cgcaaaaacgc gacgaagcgc tacatccccg gtctatatgc gcaactgtaa 240
ctaataactg atcaataattt atcattatgtt aatgcggccat taatcaagta tcgggtatata 300
acctaaggc gcaatgtattt gaatgccattt taccggccgg accgactggc ggggtgctgg 360
ggcggttaa ctgcagttat tactgcatac aagggtatca ttgcgggttat ccctgaaagg 420
taactgcagt taccacccgtt ataaatgcca tttgacgggtt gaaccgtcat gtatgttaca 480
tagtatacggg ttcatgcggg ggataactgc agttactgccc attacccggg cggaccgtaa 540
tacgggtcat gtactggat accctgaaag gatgacccgt catgttagatg cataatcagt 600
agcgataatg gtaccaatc gccaataaaacgc tcatgttagttt acccgacccat atcgccaaac 660
tgagtcccccc taaaagggtca gaggtgggtt aactgcaggat acccttcaaacc aaaaccgtgg 720
tttagttgc cctgaaaagggtt tttacagcat ttttggggcg gggtaactgc gtttacccgc 780
catccgcaca tgccaccctc cagatataattt cgtctcgaga gaccgattga tctctgggt 840
gacgaatgac cgaatagctt taattatgtt gagtatatac cctctgggtt cgaccgatcg 900
caaatttgaat ttcgaagggg cgctacgacg ggccaaaccg tgacgaggac 960
gaccggcggc cctggcgagc cccgcacccatcatgggtgac taccattacg accggacccac 1020
cgacttgggg tctaaacggta caagacacccg tctgacttgtt acgtgtactt acagggtctta 1080
cccttcaccc taatgtttagg tagtccctgtt ttttggacgtt aactatgggtt cttccgttag 1140
gacgtatcaa cggttcttca gatggggactt gacgtctatgtt gtttacccatc ttttcgttgc 1200
gttggtactt ggttaggtt gaccacgttc gccccggcg ttttgcacgtt ctgggttaggg 1260
gtgaaacact aagggtggc gacgaatcaa ccactcaaattt attcactacg ggaagagcaa 1320
ggactgttca cgtttaagaa ttttgggttcc tcctacccatc aaacgccttgg agttagaaatg 1380
accgtgtggc agcgttttctt ctgtacgtca ctcttcctcat gtttacccatc actgtatgccc 1440
tacaacgacg ggacgccttactgttcaag gctcccccatttccaaacacac aacgggtgac 1500
cgacttccctt cactgttaca ctaaagacgatc ctaccccttcc ttttgcacgtt cttacagacc 1560
accccgccctc gtctgttctt gataacgtca cccttacccatc ttttgcacgtt ttttgcacgtt 1620
ctcccttc ttcacccgtt ccaccccttctt ctttgcacgtt ttttgcacgtt ctttgcacgtt 1680
ctaccactac tccatcttcc ttttgcacgtt ctttgcacgtt ttttgcacgtt ttttgcacgtt 1740
tggtcgttaac ggtgggtgtt gtttgcacgtt ttttgcacgtt accttccatc ccaagctcaa 1800
ggatgttgc gtcgttcatg gggactacgg caactgttca tagagctctg tggaccctta 1860
ctcttacttg tacgggtaaa gtttgcacgtt ttttgcacgtt aacttccgtt ctttgcacgtt 1920
tcttacagggtt ttttgcacgtt ttttgcacgtt aacttccgtt ctttgcacgtt 1980
tttcgactat ttttgcacgtt aatgggttcaaa aagggttctt ttttgcacgtt aacacttgc 2040
tttgcgttgcgtt ttttgcacgtt ctttgcacgtt aacttccgtt ctttgcacgtt 2100
gagttactgg cggccggcggc cccggacccatc ttgtatgtt ggcgagacgtt ccgacaaggaa 2160
ggagccggag cagtgcacaa gtttgcacgtt ttttgcacgtt aacttccgtt ctttgcacgtt 2220
tcttgcgtgtt gggatttgcgtt aaggttgcgtt cacgttccatc accttgggtt ctttgcacgtt 2280
gtcttagggca ggttccatc ctgtgtggag gcacactaaa tacttgcgtt ctttgcacgtt 2340
gagggggacg agatgttgc cggacgttccatc cggccgttccatc tcaactactc 2400
gacgaatgtt ttctctttt gataacgttca ctgttccatc accttccgtt aacttccgtt 2460
tcttgcacgtt ttttgcacgtt aacttccgtt aacttccgtt ctttgcacgtt 2520
ctcgagggaaat ggcacttccatc ttttgcacgtt aacttccgtt ctttgcacgtt 2580
ccccgactgtt gacacgggtcg gtttgcacgtt ttttgcacgtt aacttccgtt ctttgcacgtt 2640
cgacggctgg ctttgcacgtt gtttgcacgtt ctttgcacgtt aacttccgtt ctttgcacgtt 2700
ctcttagagac ttttgcacgtt aacttccgtt aacttccgtt ctttgcacgtt 2760
gtatgttttta accacaagaa acgttccatc ctttgcacgtt aacttccgtt ctttgcacgtt 2820
gagtaccacc cggccacaaaca gtatgttccatc ctttgcacgtt aacttccgtt ctttgcacgtt 2880
tttcccttcat ttttgcacgtt aacttccgtt ctttgcacgtt aacttccgtt ctttgcacgtt 2940
gtcggtttttt gacgaacatc gtttgcacgtt aacttccgtt ctttgcacgtt aacttccgtt 3000
acaaagtact gtttgcacgtt aacttccgtt ctttgcacgtt aacttccgtt ctttgcacgtt 3060
tcttgcacgtt ttttgcacgtt aacttccgtt ctttgcacgtt aacttccgtt ctttgcacgtt 3120
gctcgacgtt gacgttccatc ctttgcacgtt aacttccgtt ctttgcacgtt aacttccgtt 3180
cggtcggttag acaacaaacgc gggagggggc acggaaaggaa ctggggacccatc ccacgggttag 3240
ggtgacagga aaggattttt ttttgcacgtt aacttccgtt ctttgcacgtt aacttccgtt 3300

gataagaccc cccaccccac cccgtcctgt cgttccccct cctaaccctt ctgttatcg 3360
 ccgtacgacc cctacgcccac ccgagatacc gaagactccg ctttcttgg tcgaccggca 3420
 gatccccat aggggtgcgc gggacatcgc cgcgttaattc gcgcggccca caccacaat 3480
 ggcgcgtcga ctggcgatgt gaacggtcgc gggatgcgg gcgaggaaag cgaaagaagg 3540
 gaaggaaaga gcggtcaag cggccgaaag gggcagttcg agatttagcc ccgtagggaa 3600
 atcccaaggc taaatcacga aatgccgtgg agctggggtt tttgaacta atcccaactac 3660
 caagtgcatt acccggtac gggactatct gccaaggaaagc gggaaactgc aacctcaggt 3720
 gcaagaaatt atcaccttag aacaagggtt gaccttgttg tgagttggga tagagccaga 3780
 taagaaaact aaatattccc taaaacccct aaagccggat aaccaattt ttactcgact 3840
 aaattgttt taaattgcgc ttaattaaaga cacctacac acagtcaatc ccacaccttt 3900
 caggggtccg aggggtccgt ccgtcttcat acgtttcgta cgtaggtta atcagtcgtt 3960
 ggtccacacc tttcagggtt ccgggggtc gtccgttcc atacgtttcg tacgttaggt 4020
 taatcagtcg ttgttatcg ggcggggatt gaggcgggta gggcgggat tgaggcgggt 4080
 caaggcgggt aagaggcggg gtaccgactg attaaaaaaa ataaatacgt ctccggctcc 4140
 ggcggagacg gagactcgat aaggcttca tcactcctcc gaaaaaaacct ccggatccga 4200
 aaacgtttt cgagggccct cgaacatata ggtaaaagcc tagactagtc gtgcacaact 4260
 gttaaattgtt agccgtatca tatagccgtatc tcatattatg ctgttccact ctttgattt 4320
 gtaccgggtc aactggtcac ggcaaggccca cgagttgcgc gcgtcgacg ggcctcgcca 4380
 gctcaagacc tggctggccg acaccaggcc ctgtgcact gggacaagta gtcgcggccag gtccctggtc accacggcct 4500
 gttgtgggac cggaccacca ccagcacagg tgcttgaagg cccacgcgccc ggacctgctc gacatgcggc tcaccaggct 4560
 cggcacccccc gcccctaacc gggacgcgtt gggccggccg ccctgcggag gcccggccg tactggctct agccgctcgt 4620
 gtcctcgtc ctgactgtgc acgatgctt aaagctaagg tggcggcggg agatacttc 4740
 caacccgaag ccttagcaaa aggccctgcg gccgaccctac taggaggctcg cgccccctaga 4800
 gtacgaccc aagaaggcggg tggggtttcaa aaaaataacgt cgaatattac caatgtttat 4860
 ttcgttatcg tagtggttaa agtggatttatt tcgtaaaaaa agtgcacgtaa gatcaacacc 4920
 aaacagggtt gagtagttac atagaatagt acagacatat ggcagctgga gatcgatctc 4980
 gaaccgcatt agtaccagta tcgacaaaagg acacacttta acaataggcg agtgttaagg 5040
 tgggttgtat gtcggcctt cgtatttcac atttcggacc ccacggatta ctcactcgat 5100
 tgagtgtaat taacgcacacg cgagtgcacgg gcgaaagggtc agccctttgg acagcacgg 5160
 cgacgttaatt acttagccgg ttgcgcgcct ctctccgc aacgcataac ccgcgagaag 5220
 gcaaggaggc gagtgcactga agtggatttc cgcattttatc ccaataggttgc tcttagtccc stattgcgtc ctttcttgta 5340
 cactcggttt cggcgtgtt tccggccctt ggcatttttc cggcgcaacg accgcacccg 5400
 ggtatcccgag gcggggggac ttgcgttagt gtttttagct gcgagttca gtcgcgcgc 5460
 tttgggctgt cctgatattt ctatggtccg caaaggggaa ctttcgaggg agcacggcag 5520
 aggacaaggc tgggacggcg aatggccatat ggacaggcgg aaagagggaa gcccctcgca 5580
 cccgcaaaaga gttacgatg cgacatccat agagtcaacg cacatccacg aagcgagggt 5640
 cgacccgaca cacgtgcitg gggggcaagt cgggtggcg acgcggataa ggccattgtat 5700
 agcagaactc agtttggcc attctgtgt gaatagcggtt gaccgtcgac ggtgaccatt 5760
 gtcctaatacg tctcgctcca tacatccccc acgatgtctc aagaacttca ccaccggatt 5820
 gatgccgatg tgatcttct gtcataaaacc atagacgcga gacgacttcg gtcaatggaa 5880
 gcctttttct caaccatcgaa gaactaggcc gtttgggggg tggcgaccat cgccaccaaa 5940
 aaaacaaacg ttgcgtgtct aatgcgcgtc ttttttctt agagttcttc taggaacta 6000
 gaaaagatgc cccagactgc gagtccaccc ttgttttagt gcaatccctt aaaaccatgt 6060
 ctctaataatg ttttcttaga agtggatctt gaaaaatttta atttttactt caaaatttag 6120
 ttagatttca tatatactca ttgttgcggcactt actgtcaatg ttacgaaatt agtcactccg 6180
 tggatagatg cgcttagacag ataaaagcaag taggtatcaa cggactgagg ggcagcacat 6240
 ctattgtatgc tatgcctcc cgaatggtag accgggggtca cgacgttact atggcgtct 6300
 ggggtgcgagt ggccgagggtc taaaatgtcg ttatttggtc ggtcgccctt cccggctcg 6360
 gtcttcacca ggacgttgaa ataggcggag gtaggtcaga taattaacaa cggcccttcg 6420
 atctcatca tcaagcggtc aattatcaaa cgcgttgcacaa caacgttaac gatgtccgt 6480
 gcaccacagt gcgagcagca aaccatcccg aagtaagtgc agggcaagggtt tgctagttc 6540
 cgctcaatgt actagggggtt acaacacgtt ttggccaaatc tgaggaaagc caggaggctt 6600
 gcaacagtct tcattcaacc ggcgtcacaat tagtgcgtac caataccgtc gtgacgtatt 6660
 aagagaatgtc cagtaggtca ggcattctac gaaaagacac tgaccactca tgagttgggtt 6720
 cagtagact cttatcacat acggccgttcc ctcaacgcgaa acggggccgca gttatccctt 6780
 attatggcgc ggtgtatcg tttgaaattt tcacgactgt tagtggtttttca aagaagcccc 6840
 cgcttttgcg agttcctaga atggcgacaa ctcttagtca agctacattt ggtgaggacg 6900
 tgggttgcact agaagtcgtt gaaaatgaaa gtggtcgcaaa agacccactc gtttttgtcc 6960
 ttccgtttta cggcgtttt tcccttattc ccgtgtgccc ttacaactt atgagtatga 7020

gaaggaaaaaa gttataataa cttcgtaaat agtcccaata acagagtact cgccatgt 7080
 taaacttaca taaatcttt tatttgttta tccccaaaggc gcgtgtaaag gggctttca 7140
 cggtgactg cag 7153

<210> 23
 <211> 2352
 <212> DNA
 <213> fusion protein - human

<400> 23
 atgcgtcccg gtttggact gctcctgtg gccgcctgga cggctcgccc gctggaggta 60
 cccactgatg gtaatgctgg cctgctggct gaaccccaaga ttgcctatgtt ctgtggcaga 120
 ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggacaaa 180
 acctgcattt ataccaagga aggcatctg cagtattgcc aagaagtcta ccctgaactg 240
 cagatcacca atgtgttaga agccaaaccaa ccagtgacca ttccagaactg gtgcaagcgg 300
 ggcgcgaaggc agtgcacac ccatccccac ttgtgttgc cttaccgctg cttagttgtt 360
 gagtttataa gtgtatccct tctcggttc gacaatgtca aatttttaca ccaggagagg 420
 atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagttag 480
 aagagtacca acttgcacca ctacggcatg ttgctgcctt gccaattgtt caagttccga 540
 gggtagtagt ttgtgtgtt cccactggct gaagaaaatgt acaatgttgc ttctgtgtat 600
 gcgaggaggat atgactcgga tgcgtgggtt ggcggaggcag acacagacta tgcagatggg 660
 agtgaagaca aagtatgtaga gatggcgtatg atgacgaggaa gggatggatgaggat ggaagaagaaa 720
 gaaatggatg gatggatgaggat ggtgtatgggg tagaggaaaga ggctgaggaa 780
 ccctacaaag aagccacaga gagaaccacc agcatgtcca ccaccaccc caccaccaca 840
 gagttgtgg aagagggtgt tcgatgttcc acaacacggc ccgtatcccc tgatgccgtt 900
 gacaatgtatc tcgagacacc tggggatgtt aatgaacatg cccatttcca gaaagccaaa 960
 gagaggctt agggcaagca ccggagagaga atgtcccagg tcatgagaga atgggaagag 1020
 gcagaacgtc aagcaaaagaa cttgcctaaa gctgataaga aggcatgtt ccagcatttc 1080
 caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctgggtggag 1140
 acacacatgg ccaggtgtt acgcgtatcc tggatccctc acgtgttccaa tatgtttaaag 1200
 tacatccaccc ctctgcaggc aagttatgtcc ggcgttc acgtgttccaa cggatgttccgt 1260
 ggcgttccatc ggcgttccatc aagggacaca cggccatgttccaa taaaggattt cggatgttccgt 1320
 cgcgttccatc atcccaagaa agccgtcttcc aggttatgttccaa acacccgttccgt 1380
 gtgattttatg agcgcacatgaa tcagtcttc tccctgtct acaacgttgc tgcaatgttccgt 1440
 gaggagattt cttcgatgttccaa tgatgttccaa tttcgatgttccaa agcaaaaacta ttcagatgttccgt 1500
 gtcttggcca acatgattttt tgaaccaagg atcaggtaacg gaaacgtatc tctcatgttccgt 1560
 tctttgaccg aaacgaaaac caccgtggag ctccctcccg tgaatggaga gttcagcctt 1620
 gacgatctcc agccgtgttccatc ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
 gaaggtagc ctgtgtatgc ccggccctgtt ggcgaccggag gactgaccac tcgaccatgttccgt 1740
 tctgggttcaaaatataa gacggaggag atctctgttccaa tggatgttccaa agtggatatttccgt 1800
 cgacatgact caggatataa tggatgttccaa agtggatatttccgt 1860
 ggttcaaaaca aagggtcaat cattggactc atgggtggcg gtgtgttccaa agtggatatttccgt 1920
 atcgtcatca ccttgggtat gctgaagaag aaaaagcttccgt 1980
 agtccagttt ggtggatatttccgt 2040
 tggaaaggatc caggatgttccaa gcttggatatttccgt 2100
 tgcttttccatc gccaaggatccatc ttctcatgttccaa aaggcttccgt 2160
 aaggcgaggatc agcgcacatgaa agtggatatttccgt 2220
 aaggcagatgttccatc tttcgatgttccaa agtggatatttccgt 2280
 cacctgttccatc agatgttccaa gacggatgttccaa ctttggatatttccgt 2340
 atgcagaacttccatc agtggatatttccgt 2352

<210> 24
 <211> 783
 <212> PRT
 <213> fusion protein - human

<400> 24
 Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1 5 10 15
 Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro

Gln	Ile	Ala	Met	Phe	Cys	Gly	Arg	Leu	Asn	Met	His	Met	Asn	Val	Gln
35						40					45				
Asn	Gly	Lys	Trp	Asp	Ser	Asp	Pro	Ser	Gly	Thr	Lys	Thr	Cys	Ile	Asp
50						55				60					
Thr	Lys	Glu	Gly	Ile	Leu	Gln	Tyr	Cys	Gln	Glu	Val	Tyr	Pro	Glu	Leu
65					70				75					80	
Gln	Ile	Thr	Asn	Val	Val	Glu	Ala	Asn	Gln	Pro	Val	Thr	Ile	Gln	Asn
85								90					95		
Trp	Cys	Lys	Arg	Gly	Arg	Lys	Gln	Cys	Lys	Thr	His	Pro	His	Phe	Val
100								105					110		
Ile	Pro	Tyr	Arg	Cys	Leu	Val	Gly	Glu	Phe	Ile	Ser	Asp	Ala	Leu	Leu
115								120					125		
Val	Pro	Asp	Lys	Cys	Lys	Phe	Leu	His	Gln	Glu	Arg	Met	Asp	Val	Cys
130							135				140				
Glu	Thr	His	Leu	His	Trp	His	Thr	Val	Ala	Lys	Glu	Thr	Cys	Ser	Glu
145							150				155				160
Lys	Ser	Thr	Asn	Leu	His	Asp	Tyr	Gly	Met	Leu	Leu	Pro	Cys	Gly	Ile
165								170					175		
Asp	Lys	Phe	Arg	Gly	Val	Glu	Phe	Val	Cys	Cys	Pro	Leu	Ala	Glu	Glu
180							185						190		
Ser	Asp	Asn	Val	Asp	Ser	Ala	Asp	Ala	Glu	Glu	Asp	Asp	Ser	Asp	Val
195							200						205		
Trp	Trp	Gly	Gly	Ala	Asp	Thr	Asp	Tyr	Ala	Asp	Gly	Ser	Glu	Asp	Lys
210							215					220			
Val	Val	Glu	Val	Ala	Glu	Glu	Glu	Val	Ala	Glu	Val	Glu	Glu		
225						230				235			240		
Glu	Ala	Asp	Asp	Asp	Glu	Asp	Asp	Glu	Asp	Gly	Asp	Glu	Val	Glu	Glu
245							250					255			
Glu	Ala	Glu	Glu	Pro	Tyr	Glu	Glu	Ala	Thr	Glu	Arg	Thr	Ser	Ile	
260							265					270			
Ala	Thr	Thr	Thr	Thr	Thr	Thr	Glu	Ser	Val	Glu	Glu	Val	Val	Arg	
275							280					285			
Val	Pro	Thr	Thr	Ala	Ala	Ser	Thr	Pro	Asp	Ala	Val	Lys	Tyr	Leu	
290							295					300			
Glu	Thr	Pro	Gly	Asp	Glu	Asn	Glu	His	Ala	His	Phe	Gln	Lys	Ala	Lys
305						310					315				320
Glu	Arg	Leu	Glu	Ala	Lys	His	Arg	Glu	Arg	Met	Ser	Gln	Val	Met	Arg
325							330						335		
Glu	Trp	Glu	Glu	Ala	Glu	Arg	Gln	Ala	Lys	Asn	Leu	Pro	Lys	Ala	Asp
340							345						350		
Lys	Lys	Ala	Val	Ile	Gln	His	Phe	Gln	Glu	Lys	Val	Glu	Ser	Leu	Glu
355							360						365		
Gln	Glu	Ala	Ala	Asn	Glu	Arg	Gln	Gln	Leu	Val	Glu	Thr	His	Met	Ala
370							375					380			
Arg	Val	Glu	Ala	Met	Leu	Asn	Asp	Arg	Arg	Arg	Leu	Ala	Leu	Glu	Asn
385						390				395					400
Tyr	Ile	Thr	Ala	Leu	Gln	Ala	Val	Pro	Pro	Arg	Pro	Arg	His	Val	Phe
405								410					415		
Asn	Met	Leu	Lys	Lys	Tyr	Val	Arg	Ala	Glu	Gln	Lys	Asp	Arg	Gln	His
420								425					430		
Thr	Leu	Lys	His	Phe	Glu	His	Val	Arg	Met	Val	Asp	Pro	Lys	Lys	Ala
435								440					445		
Ala	Gln	Ile	Arg	Ser	Gln	Val	Met	Thr	His	Leu	Arg	Val	Ile	Tyr	Glu
450							455					460			
Arg	Met	Asn	Gln	Ser	Leu	Ser	Leu	Leu	Tyr	Asn	Val	Pro	Ala	Val	Ala
465						470				475					480
Glu	Glu	Ile	Gln	Asp	Glu	Val	Asp	Glu	Leu	Leu	Gln	Lys	Glu	Gln	Asn
485								490					495		
Tyr	Ser	Asp	Asp	Val	Leu	Ala	Asn	Met	Ile	Ser	Glu	Pro	Arg	Ile	Ser
500								505					510		
Tyr	Gly	Asn	Asp	Ala	Leu	Met	Pro	Ser	Leu	Thr	Glu	Thr	Lys	Thr	Thr

515	520	525
Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln		
530	535	540
Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn		
545	550	555
Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr		560
565	570	575
Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser		
580	585	590
Glu Val Asn Phe Glu Val Glu Phe Arg His Asp Ser Gly Tyr Glu Val		
595	600	605
His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys		
610	615	620
Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val		
625	630	635
Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Leu Gly Thr Glu		640
645	650	655
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Leu Glu		
660	665	670
Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser Gln Pro		
675	680	685
Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe His Cys		
690	695	700
Gln Val Cys Phe Met Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys		
705	710	715
Lys Arg Arg Gln Arg Arg Ala His Gln Asn Ser Gln Thr His Gln		720
725	730	735
Ala Ser Leu Ser Lys Gln Arg Ile Ser Ser Thr Val Ala Ala Asp		
740	745	750
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn		
755	760	765
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn		
770	775	780

<210> 25
<211> 2823
<212> DNA
<213> fusion protein - human

<400> 25
atgctgcggcact gtcctgtcg gccgcctgga cggctcgccc gctggaggta 60
cccaactgatg gtaatgctgg cctgctggc gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggacaaa 180
acctgcattt ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgttgtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcacagac ccataccccac tttgtgatc cctaccgctg ctttagttgg 360
gagttataaa gtgtatgcct tctcggttcc gacaatgtca aatttataa ccaggagagg 420
atggatgtt gcgaaaactca tcttcactgg cacaccgtcg ccaaaagagac atgcagtgg 480
aagagtacca acttgcattg ctacggcatg ttgctgcctt gccggatttg caagttccga 540
ggggtagagt ttgttgttgc cccactggct gaagaaagtg acaatgtgga ttctgtctgat 600
gcggaggagg atgactcgga tgtctgtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagttagta agtagcagag gaggaagaag tggctgagggt ggaagaagaa 720
gaagccgtat atgaccggaga cgatgaggat ggtgtatggg tagaggaaga ggctgggg 780
ccctacgaag aagccacaga gagaaccacc agcatgtcca ccaccaccac caccaccaca 840
gagtctgtgg aagagggtgg tcgaggatc acaacacgcg ccagtaccac tgatgcctt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggctt aggccaagca ccgagagaga atgtccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatctt ggaacaggaa gcagccaacg agagacagca gctggtgag 1140
acacacatgg ccagagtggaa agccatgtctc aatgaccgc gccgcctggc cctggagaac 1200
tacatcaccc ctctgcaggc tggctctc cggcctcgtc acgtgttcaa tatgctaaag 1260

aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
 cgcatgggg atcccaagaa agccgctcaag atccgggtccc aggttatgac acaccccg 1380
 gtgatttatg agcgcatgaa tcagtcttc tccctgcctc acaacgtgcc tgcagtggcc 1440
 gaggagatc aggtgaagt tgatgagctg cttcagaaaag agcaaaaacta ttcatgac 1500
 gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgtgc tctcatgcca 1560
 .tctttgaccg aaacgaaaac caccgtggag ctccctcccg tgaatggaga gttcagcctg 1620
 gacgatctcc agccgtggca ttctttggg gctgactctg tgccagccaa cacagaaaac 1680
 gaagttgagc ctgttgatgc ccgcctgct gccgaccgag gactgaccac tcgaccagg 1740
 tctgggttga caaatatcaa gacggaggag atctctgaaag tgaactttga agtggaaattc 1800
 cgacatgact caggatgaa agttcatcat caagtattgg ttgttcttgc agaagatgtg 1860
 ggttcaaaa aaggtaaat cattggactc atggggcg gtgttgc atgcgacagtg 1920
 atcgtcatca ccttggatg gctgaagaag aaaaagcttgc tgaccgagctt cggatccact 1980
 agtccagttt ggttggaaatc tgcatgatata aagctatctg cttctatcgaa acaagcatgc 2040
 gatatttgc gactaaaaaa gctcaagtgc tccaaagaaaa aaccgaagtgc cgccaagtgt 2100
 ctgaagaaca actgggagtg tcgctactct cccaaaacca aaaggctcc gctgactagg 2160
 gcacatctga cagaagtggaa atcaaggctt gaaagactgg aacagctatt tctactgatt 2220
 ttccctcgag aagaccttga catgattttg aaaatggatt cttaacagga tataaaagca 2280
 ttgttaacag gattattttg acaagataat gtgaataaaag atgcccgtcac agatagattg 2340
 gettcagttt agactgatata gcctctaaca ttgagacagc atagaataaag tgcgacatca 2400
 tcatcgaaag agagtagttaa caaaaggtaa agacagttaa ctgtatcggg aattcccg 2460
 gatctggccc ccccgaccga tgcgcctg ggggacgagc tccacttaga cggcgaggac 2520
 gtggcgatgg cgcatggcga cgcgcgtac gatttcgatc tggacatgtt gggggacggg 2580
 gatccccgg gtcgggatt taccccccac gactccgccc cctacggcgc tctggatatg 2640
 gccgacttcg agttttagca gatgtttacc gatgccctt gatttgc gatcggtgg 2700
 gatatccacg acagtggcgcc cccgcacgccc gctgtcaccc cagaggagcg ccacctgtcc 2760
 aagatgcacg agaacggcta cggaaaatcca acctacaagt tctttgagca gatgcacaaac 2820
 tag 2823

<210> 26

<211> 941

<212> PRT

<213> fusion protein - human

<400> 26

Met	Leu	Pro	Gly	Leu	Ala	Leu	Leu	Leu	Ala	Ala	Trp	Thr	Ala	Arg	
1				5			10				15				
Ala	Leu	Glu	Val	Pro	Thr	Asp	Gly	Asn	Ala	Gly	Leu	Leu	Ala	Glu	Pro
				20			25				30				
Gln	Ile	Ala	Met	Phe	Cys	Gly	Arg	Leu	Asn	Met	His	Met	Asn	Val	Gln
				35			40				45				
Asn	Gly	Lys	Trp	Asp	Ser	Asp	Pro	Ser	Gly	Thr	Lys	Thr	Cys	Ile	Asp
				50			55				60				
Thr	Lys	Glu	Gly	Ile	Leu	Gln	Tyr	Cys	Gln	Glu	Val	Tyr	Pro	Glu	Leu
				65			70				75				80
Gln	Ile	Thr	Asn	Val	Val	Glu	Ala	Asn	Gln	Pro	Val	Thr	Ile	Gln	Asn
				85			90				95				
Trp	Cys	Lys	Arg	Gly	Arg	Lys	Gln	Cys	Lys	Thr	His	Pro	His	Phe	Val
				100			105				110				
Ile	Pro	Tyr	Arg	Cys	Leu	Val	Gly	Glu	Phe	Ile	Ser	Asp	Ala	Leu	Leu
				115			120				125				
Val	Pro	Asp	Lys	Cys	Lys	Phe	Leu	His	Gln	Glu	Arg	Met	Asp	Val	Cys
				130			135				140				
Glu	Thr	His	Leu	His	Trp	His	Thr	Val	Ala	Lys	Glu	Thr	Cys	Ser	Glu
				145			150				155				160
Lys	Ser	Thr	Asn	Leu	His	Asp	Tyr	Gly	Met	Leu	Leu	Pro	Cys	Gly	Ile
				165			170				175				
Asp	Lys	Phe	Arg	Gly	Val	Glu	Phe	Val	Cys	Cys	Pro	Leu	Ala	Glu	Glu
				180			185				190				
Ser	Asp	Asn	Val	Asp	Ser	Ala	Asp	Ala	Glu	Glu	Asp	Asp	Ser	Asp	Val
				195			200				205				
Trp	Trp	Gly	Gly	Ala	Asp	Thr	Asp	Tyr	Ala	Asp	Gly	Ser	Glu	Asp	Lys
				210			215				220				

Val Val Glu Val Ala Glu Glu Glu Val Ala Glu Val Glu Glu Glu
 225 230 235 240
 Glu Ala Asp Asp Asp Glu Asp Asp Gly Asp Glu Val Glu Glu
 245 250 255
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
 260 265 270
 Ala Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
 275 280 285
 Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
 290 295 300
 Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
 305 310 315 320
 Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
 325 330 335
 Glu Trp Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
 340 345 350
 Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
 355 360 365
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
 370 375 380
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
 385 390 395 400
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
 405 410 415
 Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
 420 425 430
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
 435 440 445
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
 450 455 460
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
 465 470 475 480
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
 485 490 495
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
 500 505 510
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
 515 520 525
 Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
 530 535 540
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
 545 550 555 560
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
 565 570 575
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
 580 585 590
 Glu Val Asn Phe Glu Val Glu Phe Arg His Asp Ser Gly Tyr Glu Val
 595 600 605
 His His Gln Val Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
 610 615 620
 Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
 625 630 635 640
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
 645 650 655
 Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
 660 665 670
 Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
 675 680 685
 Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
 690 695 700
 Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
 705 710 715 720

Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
 725 730 735
 Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
 740 745 750
 Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
 755 760 765
 Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
 770 775 780
 Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
 785 790 795 800
 Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
 805 810 815
 Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp
 820 825 830
 Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala
 835 840 845
 Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly
 850 855 860
 Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met
 865 870 875 880
 Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp
 885 890 895
 Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Ala Asp Ala Ala
 900 905 910
 Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr
 915 920 925
 Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
 930 935 940

<210> 27
 <211> 5015
 <212> DNA
 <213> human

<400> 27
 gacggatcgg gagatctccc gatccccat ggtcgactct cagtacaatc tgctctgatg 60
 ccgcatagtt aagccagtat ctgctccctg ctttgtgtt ggaggctcgct gagtagtgcg 120
 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatt aagaatctgc 180
 tttagggttag gcgttttgcg ctgcattcgc atgtacgggc cagatatacg cgttgacatt 240
 gattatttgcg tagtattaa tagaatcaa ttacgggttc attagttcat agcccatata 300
 tggagttccg cgttacataa cttacggtaa atggccgccc tggctgaccg cccaaacgacc 360
 ccccccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
 attgacgtca atgggtggac tatttacggt aaactgccc cttggcagta catcaagtgt 480
 atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
 atgcccagta catgaccta tgggacttc ctacttggca gtacatctac gtattagtca 600
 tcgcttattac catggtgatg cggttttggc agtacatcaa tgggcgtgg tagcggtttg 660
 actcacgggg atttccaagt ctccacccca ttgacgtca tgggagttt ttttggcacc 720
 aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatggcg 780
 gtaggcgtgt acggtgggag gtctataaa gcagagctct ctggcttaact agagaaccca 840
 ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc 900
 gtttaaactt aagcttggta ccgagctcg atccactagt ccagtgtggt ggaattctgc 960
 agatattccag cacagtggcg gccgctcgag tctagagggc ccgtttaaac ccgctgatca 1020
 gcctcgactg tgccttctag ttgcccggca tctgttggc gcccctcccc cgtgccttcc 1080
 ttgaccctgg aagggtggcac tcccactgtc ctttccataat aaaaatgggaga aattgcatcg 1140
 cattgtctga gttaggtgtca ttcttattctg ggggggggg tggggcagga cagcaagggg 1200
 gaggattggg aagacaatacg caggcatgt ggggatgcgg tgggctctat ggcttctgag 1260
 gcgaaaagaa ccagctgggg ctctaggggg tatccccacg cggccctgttag cggcgcatta 1320
 agcgcggcgg gtgtgggt tacgcgcagc gtgaccgcta cacttgcacag cggccctagcg 1380
 cccgctccct tcgcttctt cccttcctt ctcgcccacgt tcgcccggctt tccccgtcaa 1440
 gctctaaatc gggcatccc tttaggggtc cgatttagtg cttagccgca cctcgaccgg 1500
 aaaaaacttg attagggtga tggttcacgt agtggccat cggccctgata gacgggtttt 1560

cgccctttga	cgttggagtc	cacgttctt	aatagtggac	tcttgggcc	aactggaaaca	1620
acactcaacc	ctatccgggt	ctattcttt	gatttataag	ggattttggg	gatttcggcc	1680
tattggtaa	aaaatgagct	gatttaacaa	aaatttaacg	cgaattaatt	ctgtggaaatg	1740
tgtgtcagg	agggtgtgga	aagtccccag	gctcccccagg	waggcagaag	tatgcaaagc	1800
atgcatctca	attagtcagc	aaccaggtgt	gaaaaatccc	caggctcccc	agcaggcaga	1860
agtatgcaaa	gcatgcatct	caattagtc	gcaaccatag	tcccggccct	aactccgccc	1920
atcccgcccc	taatccggcc	cagttccggc	catttccgc	ccatggctgc	actaatttt	1980
tttattttatg	cagaggccga	ggccggctct	gctcttgagc	tattccgaaa	gtatgtgagg	2040
ggcttttttg	gaggctttag	cttttgc当地	aagctccccgg	gagcttgtat	atccatttt	2100
ggatctgatc	agcacgtgtt	gacaattaaat	catcgccata	gtatatcgcc	atagtataat	2160
acgacaaggt	gaggaactaa	accatggcca	agttgaccag	tgccgttccg	gtgctaccg	2220
cgcgcgacgt	cgcggagcg	gtcgaggctt	ggaccgaccg	gctcgggttc	tcccggact	2280
tcgtggagga	cgacttcgccc	ggtgtggccc	gggacgacgt	gaccctgttc	atcagcgcgg	2340
tccaggacca	ggtgtgtccg	gacaacaccc	tggccctgggt	gtgggtgcgc	ggcctggacg	2400
agctgtacgc	cgagtgtcg	gaggtctgt	ccacacactt	ccggggacgc	tccggggccgg	2460
ccatgaccca	gatccggcag	cagccctggg	ggcggggagg	cgcggctgcgc	ggccggccgg	2520
gcaatgcgt	gcacttcgt	gcccggaggc	agactgaca	cgtgtctacga	gatttcgatt	2580
ccaccggccgc	cttctatgaa	aggttgggtt	tcggaaatcg	tttccgggac	ggccggctgg	2640
tgatccttca	gcccggggat	ctcatgtgg	agtttcttcgc	ccaccccaac	ttgttattt	2700
cagcttataa	tggttacaaa	taaagcaata	gcatcacaat	tttcacaaat	aaagcattt	2760
tttcaactgca	ttctatgtt	gggttgc当地	aactcatcaa	tgatctttat	catgtctgt	2820
taccgtcgac	ctctatgtt	agcttggcgt	aatcatggtc	atagctgtt	cctgtgtgaa	2880
attgttatacc	gctcacaatt	ccacacaca	tacgagccgg	aagcataaaag	tgtaaagcc	2940
gggggtgccta	atagtgagc	taactcacat	taatttgcgtt	ggcgtctactg	cccgcttcc	3000
agtcggggaaa	cctgtctgtc	cagctgcatt	aatgaatccg	ccaacgcgcg	ggggagaggcg	3060
gtttgcgtat	tgggcgctct	tccgcttct	cgctcaactga	ctcgctgcgc	tcggctgttc	3120
ggctgcggcg	agcggatata	gctcactcaa	aggcggtat	acggttatcc	acagaatcag	3180
gggataaacgc	agggaaaagaa	atgtgagcaa	aaggccagca	aaaggccagg	aaccgtaaaa	3240
aggccgcgtt	gctggcggtt	ttccatagc	tccggcccccc	tgacgagcat	cacaaaaatc	3300
gacgctcaag	tcagagggtg	cgaaaacccg	caggactata	aatgataccg	ggctttccccc	3360
ctggaaatgc	cctctgtgc	tctctgtc	cgaccctgc	gcttacccgg	tacctgtccg	3420
ccttcttccc	tccggaaatg	gttgcgtt	ctcaatgtc	acgctgttag	tatctctagg	3480
cgtgttaggt	cgttcgtcc	aagctgggt	gtgtgcacga	accccccgtt	cagcccgacc	3540
gctgcgcctt	atccggtaac	tatctgtctt	agtccaaatcc	ggtaagacac	gacttatacg	3600
cactggcagc	agccacttgg	aacaggatta	gcagagcgg	gtatgttagg	ggtgctacag	3660
agttcttga	gtggggcct	aactacggc	acactagaag	gacagtattt	ggtatctcg	3720
ctctgtctaa	gccaggattc	ttcggaaaaaa	gagtttggtag	ctcttgatcc	ggcaaaacaaa	3780
ccaccgcgtt	tagccgtgg	ttttttgttt	gcaaggcagca	gattacgcgc	agaaaaaaaaag	3840
gatctcaaga	agatcccttt	atctttctt	cggggctctg	cgctctgtt	aaacggaaaact	3900
cacgtttaagg	gattttggc	atagattat	aaaaaggat	tcttcacat	atccctttaa	3960
attaaaaatg	aaatgtttaa	tcaatctaa	gtatataatg	gtaaaacttgg	tctgacagt	4020
accaatgttt	aatcagtgt	gcacctatct	cagcgatct	tctatttcgt	tcatccatag	4080
ttgcctgtact	ccccgtcg	tagataacta	cgatacggg	gggcttacca	tctggccccca	4140
gtgctgcaat	gataccgcg	gaccggacgt	caccggctcc	agattttatc	gcaataaaacc	4200
agccagccgg	aaggcccgag	cgcagaatgt	gtcctgtcaac	tttatccgc	tccatccagt	4260
ctattaattt	ttggccggaa	gctagagttt	gtatgtcgcc	agttatagt	ttgcgcacg	4320
tttgtgcata	tgctacaggc	atctgtgggt	cacgtctgtc	gttgggtat	gtttcattca	4380
gctccgggtt	ccaaatgcgt	aggcgaggta	catgtcccc	catgttgc	aaaaaaagcgg	4440
tttagctctt	ccgttctccg	atctgtgtc	gaagtaatgt	ggccgcagtg	ttatctactc	4500
tggttatggc	agcactgtat	aattcttctt	ctgtcatgcc	atccgtaaaga	tgctttctg	4560
tgactgtgt	gtactcaacc	aagtcttctt	gagaatagt	tatgcggcga	ccgagttgct	4620
cttgcggcgc	gtcaatacgg	gataataccg	cgccacatag	cagaacttta	aaagtgtca	4680
tcatttggaaa	acgttcttcg	gggcggaaaac	tctcaaggat	cttaccgcgt	ttgagatcca	4740
gttcgtatgt	acccactcg	gcacccaact	gatcttcagc	atcttttact	ttcaccagcg	4800
tttctgggt	agcaaaaaaa	ggaaggccaa	atgcggccaa	aaagggaata	aggcgcacac	4860
ggaaatgttg	aataactcata	ctcttcctt	tcaatattt	ttgaagcatt	tatcagggtt	4920
attgtctcat	gaggcgatc	atatttgaat	gtatttagaa	aaataaaacaa	atagggttc	4980
cgccgcacatt	tcccccggaaa	gtgcccaccc	acgtc			5015

<211> 5015
<212> DNA
<213> human

<400> 28

ctgcctagcc	ctctagaggg	ctaggggata	ccagctgaga	gtcatgttag	acgagactac	60
ggcgatataaa	ttcggtcata	gacgaggagac	gaacacacaa	cctccagcga	ctcatcacgc	120
gctcgttta	aattcgatgt	tgttccgtc	cgaactggct	gttaacgtac	ttcttagacg	180
aatccccatc	cgaaaaacgc	gacgaagcgc	tacatccccg	gtctatatgc	gcaactgtaa	240
ctaataactg	atcaataatt	atcattagtt	aatgcggccag	taatcaagta	tcgggtat	300
acctaaggc	gcaatgtatt	gaatgccatt	taccggccgg	accgactggc	gggttgcgtgg	360
gggcgggtaa	ctgcagttat	tactgcatac	aagggtatca	ttgcggtat	ccctgaaagg	420
taactgcgt	taccacctg	ataaatgcct	tttgcgggtt	gaaaccgtat	gtagttcaca	480
tagtatacgg	ttcatgcggg	ggataactgc	agttactgc	atttacccggg	cgaccgtaa	540
tacgggtcat	gtactggaat	accctgaaa	gatgaaccgt	catgttagat	cataatcagt	600
agcgataatg	gtaccactac	gccaaaaccc	tcatgttagt	acccgcac	atcgccaaac	660
tgagtgc	taaaggttca	gagggtgggt	aactgcagtt	acccctcaaa	aaaaccgtgg	720
ttttagtgtc	cctgaaaagg	tttacagcat	tgttgcggcg	ggtaactgc	gtttacccgc	780
catccgcaca	tgccacccct	catatatatt	cgtctcgaga	gaccgattga	tctcttgggt	840
gacgaatgac	cgaatagctt	taattatgt	gagtgtat	cctctgggtt	cgaccgtatc	900
caaatttggaa	ttcgaaccat	ggctcgagcc	taggtgtat	gttcacacca	ccttaagacg	960
tctataggtc	gtgtcaccgc	cgccgagc	agatctccc	ggcaatttgc	ggcgactagt	1020
cgaggctgac	acggaagatc	aacggtcgtt	agacaacaaa	cggggagggg	gcacgaaagg	1080
aactgggacc	ttccacgg	agggtgacag	gaaaggattt	tttactcct	ttaacgtac	1140
gtaacagact	catccacagt	aagataagac	ccccccaccc	accccgctt	gtcgcccc	1200
ctcctaacc	ttctgttata	gtccgtac	ccccctacgc	acccgagata	ccgaagactc	1260
cgccttctt	ggtcgacccc	gagatcccc	atagggtgc	gggggacatc	gccgcgtat	1320
tcgcgcgc	cacaccacca	atgcgcgtc	cactggcgt	gtgaacggc	gccccatc	1380
ggcgagggaa	agcggaaagaa	gggaaggaaa	gagcgggtca	acccggccaa	aggggcagtt	1440
cgagatttag	ccccgtaggg	aaatcccaag	gctaatac	gaaatgcgt	ggagctgggg	1500
tttttgaac	taatcccact	accaagtgc	tcacccggta	gggggactat	ctgcca	1560
gcgggaaact	gcaacctcag	gtgcaagaaa	ttatcacctg	agaacaagg	ttgaccc	1620
tgtgagtgg	gataagggca	gataagaaaa	ctaaattatc	ctaaaaaccc	ctaaaggccg	1680
ataaccaatt	tttactcga	ctaaatttt	tttaatttgc	gottaattaa	gacaccc	1740
acacagtca	ttccacac	ttcagggtt	cgaggggtt	gtccgtt	atacgtt	1800
tactgttagt	taatccatc	tttgttccaca	ccttccagg	gtccgagg	tcgtccgt	1860
tcatacgtt	cgtacgtaga	gttaatcgt	cgttgtatc	agggcg	ttgaggccg	1920
tagggccgg	attgaggccg	gtcaaggccg	gtaagaggcg	gggtaccgac	tgattaaaa	1980
aaataaaatac	gtctccgg	ccggcggaga	cgagactcg	ataagg	catcact	2040
ccgaaaaaac	ctccggatcc	aaaaacgtt	ttcggggcc	ctcgaacata	tagtaaa	2100
ccttagact	tcgtgcacaa	ctgttaat	gtagccgtat	catataggc	tatcatat	2160
tgctgttcca	ctccgtt	ttgttacgg	tcaactgg	acggcaagg	cacgatgg	2220
gcgcgctgca	gcccgc	cagctcaaga	cctggcgg	cgagccaa	agggcc	2280
agcacctt	gctgaagccg	ccacaccagg	ccctgtc	ctgggaca	tagtgc	2340
aggtccgtt	ccaccacggc	ctgttgtgg	accggacca	cacccacgc	ccggac	2400
tcgacatgc	gctcaccagg	ctccagcaca	ggtgc	ggccctgc	aggcccgg	2460
ggtaactgg	ctaggcgtc	gtcggcac	ccgcctc	gggggacgc	ctggccgg	2520
cgttgacca	cgttgcac	cggtccctcg	tcctgtact	gacatgt	ctaaagct	2580
ggtggccgg	aaagatactt	tccaaaccc	agccttgc	aaaggccct	cgccgac	2640
actaggagg	cgcgcctt	gatgtac	tcaagaagcg	ggtgggtt	aacaaataa	2700
gtcgaatatt	accaatgtt	atttgcgtt	cgtgtt	aaagtgtt	ttcgta	2760
aaagtgcgt	aaatcaaca	ccaaacagg	ttgatgtt	acatagaata	gtacagac	2820
atggcagtc	gagatc	tcaacc	ttagtacc	tatcgac	ggacacact	2880
taacaatagg	cgatgtt	ggtgtt	atgcctgg	ttcgat	acatttgc	2940
ccccacggat	tactcact	attgtat	attaacgc	cgcgatgc	gggcgaaagg	3000
tcagcccc	ggacac	gtcgacgt	ttacttag	ggttgc	ccctccgc	3060
caaaccgcata	acccgc	aggcga	gacgt	gacgcac	agccagca	3120
ccgacccgc	tcgc	aggcgtt	tccgcatt	tccaaat	tgtcttgc	3180
ccctatttgc	tccttctt	tacactc	ttccggcgt	ttccgg	ttggcattt	3240
tccggccaa	cgacc	aagg	aggcgggg	actgc	gtgttt	3300
ctgcgagg	agtcc	gttgg	gtcctgat	ttctatgg	cgcaaagg	3360
gaccc	ggac	acgc	agaggaca	gtgg	atggacagg	3420

ggaaagaggg	aaggcccttcg	caccgcgaaa	gagttacgag	tgcgacatcc	atagagtcaa	34800
gccacatcca	gcaagcggagg	ttcgaccggc	cacacgtgt	tggggggca	gtcgggctgg	3540
cgacgcgaa	taggccattg	atagcagaac	tcaggttggg	ccattctgt	ctgaatagcg	36000
gtgaccgtcg	tcggtgacca	ttgtccta	cgtctcgctc	catacatccg	ccacgatgtc	3660
tcaagaacctt	caccacgg	ttgatgccc	tgtgatcttc	ctgtcataaa	ccatagacgc	3720
gagacgactt	cgggtcaatgg	aagcctttt	ctcaaccatc	gagaactagg	ccgtttttt	3780
ggtgtccgacc	atcgccacca	aaaaaacaaa	cgttgcgt	ctaatgcgc	tcttttttc	3840
cttagtgtt	tcttagaaac	tagaaaaatg	gccccagact	gcgagtcacc	ttgcttttga	39000
gtgcaatcc	ctaaaaccag	tactctaata	gtttttctta	gaagtggatc	tagaaaaatt	3960
taatttttac	ttcaaaaattt	agtttagattt	catatatact	catttgaacc	agactgtcaa	4020
tggttacgaa	ttagtcaactc	cgtggataga	gtcgctagac	agataaaagca	agtaggtatc	4080
aacggactga	ggggcagcac	atctattgtat	gctatgcct	cccgaatgg	agaccggggt	4140
cacgacgtt	ctatggcgct	ctgggtgcga	gtggccgagg	tctaaatagt	cgttattttgg	4200
tcggtcgccc	ttcccggtc	gcgttctcac	caggacgttg	aaataggcgg	aggtaggtca	4260
gataattaac	aacggccctt	cgtatccatt	catcaagcgg	tcaattatca	aacgcgttgc	4320
aacaacggtt	acgtatgtccg	tagcaccaca	gtgcgagcag	caaacatata	cgaagtaagt	4380
caggccaaag	ggttgtcttagt	tccgtctcaat	gtactagggg	gtacaacacg	tttttcgcc	4440
aatcgagggaa	ggcaggaggc	tagcaacagt	cttcattcaa	ccggcgtcac	aatagtgtat	4500
accaataccg	tcgtgacgta	ttaagagaat	gacagtacgg	taggcattct	acgaaaaagac	4560
actgaccact	catgagttgg	ttcagtaaga	ctcttatcac	atacgcccgt	ggctcaacga	4620
gaacggggcg	cagttatgcc	ctattatggc	gcgggtgtatc	gtcttgaat	tttcacgagt	4680
agtaacccctt	tgcaagaagc	cccgctttt	agagttccta	gaatggcgcac	aactcttagt	4740
caagctacat	tgggtgagca	cgtgggttga	ctagaagtcg	tagaaaaatg	aagtggtcgc	4800
aaagacccac	tcgtttttgt	ccttccgtt	tacggcgtt	tttcccttat	cccgctgtg	4860
cctttacac	ttatgatgt	gagaaggaaa	aagttaataat	aacttcgtaa	atagtcccaa	4920
taacagagta	ctcgccctatg	tataaaactt	cataaatctt	tttatttgtt	tatccccaaag	4980
qcgcqgtqaa	aggggctttt	cacggggac	tgcag			5015

<210> 29
<211> 2352
<212> DNA
<213> fusion protein - human

tctttgaccg aaacgaaaac caccgtggag ctcctcccg tgaatggaga gttcaggctg 1620
 gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
 gaagttgagc ctgttgatgc ccggccctgt gccgaccgag gactgaccac tcgaccagg 1740
 tctgggttga caaatatcaa gacggaggag atctctgaag tgaactttga agtggaaattc 1800
 cgacatgact caggatatga agttcatcat caaaaattgg ttttctttgc agaagatgtg 1860
 ggtaaaca aaggtgcaat cattggactc atggggcg gtgttgtcat agcgacagt 1920
 atcgcatca ccttggatgat gctgaagaag aaaaagctt gtaccgagct cgatccact 1980
 agtccagttt ggtggaaattc tgcagatatac ctggagccag tagatccttag actagagccc 2040
 tggaaagcatc caggaagtca gcctaaaact gcttgtacca attgctattt taaaaagtgt 2100
 tgcttcatt gccaagttt tttcatgaca aaagcttag gcatctccta tggcaggaaag 2160
 aagcggagac agcgcacgaag agtcatcag aacagtca ctcataaagc ttctctatca 2220
 aagcagagga tatccagcac agtggccggc gcagacgccc ctgtcacc 2280
 cacctgtcca agatgcagca gaacggctac gaaaatccaa cttacaagtt ctttggcag 2340
 atgcagaact ag 2352

<210> 30

<211> 783

<212> PRT

<213> fusion protein - human

<400> 30

Met	Leu	Pro	Gly	Leu	Ala	Leu	Leu	Leu	Ala	Ala	Trp	Thr	Ala	Arg	
1	5						10				15				
Ala	Leu	Glu	Val	Pro	Thr	Asp	Gly	Asn	Ala	Gly	Leu	Leu	Ala	Glu	Pro
							20				25			30	
Gln	Ile	Ala	Met	Phe	Cys	Gly	Arg	Leu	Asn	Met	His	Met	Asn	Val	Gln
							35				40			45	
Asn	Gly	Lys	Trp	Asp	Ser	Asp	Pro	Ser	Gly	Thr	Lys	Thr	Cys	Ile	Asp
							50				55			60	
Thr	Lys	Glu	Gly	Ile	Leu	Gln	Tyr	Cys	Gln	Glu	Val	Tyr	Pro	Glu	Leu
							65				70			75	
															80
Gln	Ile	Thr	Asn	Val	Val	Glu	Ala	Asn	Gln	Pro	Val	Thr	Ile	Gln	Asn
							85				90			95	
Trp	Cys	Lys	Arg	Gly	Arg	Lys	Gln	Cys	Lys	Thr	His	Pro	His	Phe	Val
							100				105			110	
Ile	Pro	Tyr	Arg	Cys	Leu	Val	Gly	Glu	Phe	Ile	Ser	Asp	Ala	Leu	Leu
							115				120			125	
Val	Pro	Asp	Lys	Cys	Lys	Phe	Leu	His	Gln	Glu	Arg	Met	Asp	Val	Cys
							130				135			140	
Glu	Thr	His	Leu	His	Trp	His	Thr	Val	Ala	Lys	Glu	Thr	Cys	Ser	Glu
							145				150			155	
															160
Lys	Ser	Thr	Asn	Leu	His	Asp	Tyr	Gly	Met	Leu	Leu	Pro	Cys	Gly	Ile
							165				170			175	
Asp	Lys	Phe	Arg	Gly	Val	Glu	Phe	Val	Cys	Cys	Pro	Leu	Ala	Glu	Glu
							180				185			190	
Ser	Asp	Asn	Val	Asp	Ser	Ala	Asp	Ala	Glu	Glu	Asp	Asp	Ser	Asp	Val
							195				200			205	
Trp	Trp	Gly	Gly	Ala	Asp	Thr	Asp	Tyr	Ala	Asp	Gly	Ser	Glu	Asp	Lys
							210				215			220	
Val	Val	Glu	Val	Ala	Glu	Glu	Glu	Val	Ala	Glu	Val	Glu	Glu		
							225				230			235	
Glu	Ala	Asp	Asp	Glu	Asp	Asp	Glu	Asp	Gly	Asp	Glu	Val	Glu	Glu	
														240	
Glu	Ala	Glu	Glu	Pro	Tyr	Glu	Glu	Ala	Thr	Glu	Arg	Thr	Thr	Ser	Ile
							245				250			255	
Ala	Thr	Thr	Thr	Thr	Thr	Glu	Ser	Val	Glu	Glu	Val	Glu	Glu		
							260				265			270	
Val	Pro	Thr	Thr	Ala	Ala	Ser	Thr	Pro	Asp	Ala	Val	Asp	Lys	Tyr	Leu
							275				280			285	
Glu	Thr	Pro	Gly	Asp	Glu	Asn	Glu	His	Ala	His	Phe	Gln	Lys	Ala	Lys
							290				295			300	
Glu	Arg	Leu	Glu	Ala	Lys	His	Arg	Glu	Arg	Met	Ser	Gln	Val	Met	Arg
															320

	325	330	335												
Glu	Trp	Glu	Glu	Ala	Glu	Arg	Gln	Ala	Lys	Asn	Leu	Pro	Lys	Ala	Asp
		340	345	350											
Lys	Lys	Ala	Val	Ile	Gln	His	Phe	Gln	Glu	Lys	Val	Glu	Ser	Leu	Glu
		355	360	365											
Gln	Glu	Ala	Ala	Asn	Glu	Arg	Gln	Gln	Leu	Val	Glu	Thr	His	Met	Ala
		370	375	380											
Arg	Val	Glu	Ala	Met	Leu	Asn	Asp	Arg	Arg	Arg	Leu	Ala	Leu	Glu	Asn
		385	390	395											400
Tyr	Ile	Thr	Ala	Leu	Gln	Ala	Val	Pro	Pro	Arg	Pro	Arg	His	Val	Phe
		405	410	415											
Asn	Met	Leu	Lys	Tyr	Val	Arg	Ala	Glu	Gln	Lys	Asp	Arg	Gln	His	
		420	425	430											
Thr	Leu	Lys	His	Phe	Glu	His	Val	Arg	Met	Val	Asp	Pro	Lys	Lys	Ala
		435	440	445											
Ala	Gln	Ile	Arg	Ser	Gln	Val	Met	Thr	His	Leu	Arg	Val	Ile	Tyr	Glu
		450	455	460											
Arg	Met	Asn	Gln	Ser	Leu	Ser	Leu	Leu	Tyr	Asn	Val	Pro	Ala	Val	Ala
		465	470	475											480
Glu	Glu	Ile	Gln	Asp	Glu	Val	Asp	Glu	Leu	Leu	Gln	Lys	Glu	Gln	Asn
		485	490	495											
Tyr	Ser	Asp	Asp	Val	Leu	Ala	Asn	Met	Ile	Ser	Glu	Pro	Arg	Ile	Ser
		500	505	510											
Tyr	Gly	Asn	Asp	Ala	Leu	Met	Pro	Ser	Leu	Thr	Glu	Thr	Lys	Thr	Thr
		515	520	525											
Val	Glu	Leu	Leu	Pro	Val	Asn	Gly	Glu	Phe	Ser	Leu	Asp	Asp	Leu	Gln
		530	535	540											
Pro	Trp	His	Ser	Phe	Gly	Ala	Asp	Ser	Val	Pro	Ala	Asn	Thr	Glu	Asn
		545	550	555											560
Glu	Val	Glu	Pro	Val	Asp	Ala	Arg	Pro	Ala	Ala	Asp	Arg	Gly	Leu	Thr
		565	570	575											
Thr	Arg	Pro	Gly	Ser	Gly	Leu	Thr	Asn	Ile	Lys	Thr	Glu	Glu	Ile	Ser
		580	585	590											
Glu	Val	Asn	Phe	Glu	Val	Glu	Phe	Arg	His	Asp	Ser	Gly	Tyr	Glu	Val
		595	600	605											
His	His	Gln	Lys	Leu	Val	Phe	Phe	Ala	Glu	Asp	Val	Gly	Ser	Asn	Lys
		610	615	620											
Gly	Ala	Ile	Ile	Gly	Leu	Met	Val	Gly	Gly	Val	Val	Ile	Ala	Thr	Val
		625	630	635											640
Ile	Val	Ile	Thr	Leu	Val	Met	Leu	Lys	Lys	Lys	Leu	Gly	Thr	Glu	
		645	650	655											
Leu	Gly	Ser	Thr	Ser	Pro	Val	Trp	Trp	Asn	Ser	Ala	Asp	Ile	Leu	Glu
		660	665	670											
Pro	Val	Asp	Pro	Arg	Leu	Glu	Pro	Trp	Lys	His	Pro	Gly	Ser	Gln	Pro
		675	680	685											
Lys	Thr	Ala	Cys	Thr	Asn	Cys	Tyr	Cys	Lys	Cys	Cys	Phe	His	Cys	
		690	695	700											
Gln	Val	Cys	Phe	Met	Thr	Lys	Ala	Leu	Gly	Ile	Ser	Tyr	Gly	Arg	Lys
		705	710	715											720
Lys	Arg	Arg	Gln	Arg	Arg	Arg	Ala	His	Gln	Asn	Ser	Gln	Thr	His	Gln
		725	730	735											
Ala	Ser	Leu	Ser	Lys	Gln	Arg	Ile	Ser	Ser	Thr	Val	Ala	Ala	Asp	
		740	745	750											
Ala	Ala	Val	Thr	Pro	Glu	Glu	Arg	His	Leu	Ser	Lys	Met	Gln	Gln	Asn
		755	760	765											
Gly	Tyr	Glu	Asn	Pro	Thr	Tyr	Lys	Phe	Phe	Glu	Gln	Met	Gln	Asn	
		770	775	780											

<210> 31
<211> 2823
<212> DNA

<213> fusion protein - human

<400> 31

atgctgcccg gttggcaact gtcctgtc gccccttga cggctcccc gctggaggta 60
 cccactgatg gtaatgtcgg cctgtggct gaacccaga ttgcattgt ctgtggcaga 120
 ctgaacatgc acatgaatgt ccagaatggg aagtggatt cagatccatc agggacccaa 180
 acctgcattg ataccaagga aggcatctg cagtattgcc aagaagtcta ccctgaactg 240
 cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
 ggcccaagc agtgcacagac ccatccccac tttgtgattc cctaccgctg cttagtttgt 360
 gagttataa gtatgtccct tctcggttccct gacaatgtca aattcttaca ccaggagagg 420
 atggatgttt gcgaaaactca tcttcacttg cacaccgtcg ccaaagagac atgcagtgtg 480
 aagagtacca acttgcatga tcacggcatg ttgtgtccct gcggaaatgtga caagttccga 540
 ggggtagagt ttgtgtgttgc cccactggct gaagaaatgtg acaatgtggta ttctgctgtat 600
 gcggaggagg atgactcgga tgcgtgtgttgc ggcggagcag acacagacta tgcagatggg 660
 agtgaagaca aagtagttaga agtagcagag gaggaagaag tggctgagggt ggaagaagaa 720
 gaagccatgt atgacgagga cgatgaggat ggtgtatgagg tagaggaaga ggctgaggaa 780
 ccctacgaag aagccacaga gagaaccacc accattgtcca ccaccaccac caccaccaca 840
 gagtctgtgg aaggagggtgt tcgcgttccct acaacacgac ccagtaccct tgatgccgtt 900
 gacaatgtc tcgagacacc tggggatgatg aatgacatcg cccatgttcca gaaagccaa 960
 gagaggcttgc agggcaagca cccggagagaa atgtcccttgc tcatgagaga atgggaagag 1020
 gcagaacgtc aagccaaagaa cttgcctaa gctgataaga aggcagttat ccagcatttc 1080
 caggagaaaag tggaaatctt ggaacaggaa gcagccaaacg agagacagca gctgggtggag 1140
 acacacatgg ccagactgttgc agccatgtc aatgaccgttgc gccccttgc cctggagaac 1200
 tacatcaccgc ctctgcaggc tgggttccct cggctcgatc acgtgttca tatgctaaag 1260
 aagtatgtcc gcgcagaaca gaaggacaga cagcacacccc taaaaggatcgatc cgagcatgtg 1320
 cgcgtgtgg atcccaagaa agccgtctcg atccggctcc agtttatgtac acacccgt 1380
 tgattttatgc agcgtatggatc tcgtctc tccctgttca acaacgtgtcc tgcgtggcc 1440
 gaggagattc aggatgttgc tgatgatgtc cttcagaaaag agccaaacta ttcagatgtac 1500
 gtcttggcca acatgatttgc tgaaccaagg atcagttacg gaaacgtgc tctcatgtcca 1560
 tctttgaccgc aaacgaaaac caccgtggag ctccctcccg tgaatggaga gttcagctg 1620
 gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
 gaagttgagc ctgttgcattgc ccggccctgt gccgaccggag gactgaccac tcgaccagg 1740
 tctgggttgc caaatatgttgc gacggggag atctgttgc tgaatgttgc agtggatttgc 1800
 cgacatgtactt caggatgttgc attcatcatc caaaaattgg tttcttttgc agaagatgtg 1860
 ggttcaaaaca aagggtcaat cattggactc atgggtggcc gttgtgtcat agcgcacatgt 1920
 atcgtcatca ccttgggtatgc gctgaaagaa aaaaagcttgc gtaccgagct cggatccact 1980
 agtccagttgt ggttgcatttgc tgacatgttgc aagctactgt cttctatgtca acaagcatgtc 2040
 gatatttgc gacttaaaaaa gctcaagtgc tccaaagaaa aaccgaagtgc cgccaaatgt 2100
 ctgaagaaaca actggggatgt tcgctactt cccaaacca aaaggctcc gctgactagg 2160
 gcacatgtca cagaagtggaa atcaaggcttca gaaagactgg aacagcttatt tctactgtt 2220
 tttcctcgatgg aagaccttgc catgatggatc aaaatggatttgc ttttacaggaa tataaaagca 2280
 ttgtttaacag gatttttgc tcaagataat gtgataaaatgc atggcgatc agatagatttgc 2340
 gcttcagtgg agactgtatgc gcctctaaaca ttgagacagc atagaataag tgcgacatca 2400
 tcacggaaag agagtagttaa caaaaggcttca agacagttgc ctgtatcggtt aattcccggt 2460
 gatctggccc cccccggccca tgcgtatcgatc ggggacggc tccacttgc tggcggatgt 2520
 gtggcgatgg cgcgtatcgatc cgcgtatcgatc gatggatgttgc tggcggatgttgc 2580
 gatccggccgg gtcggggatgg taccggccac gactggccccc cctacggccgc tctggatgt 2640
 gcccggatgttgc agtttgcattgc gatgttttgc gatggcccttgc gaaatgttgc gatggatgttgc 2700
 gatatccacgc acatggccgg ccggcgtatcgatc gatggatgttgc cccacccgttgc 2760
 aagatgcacgc agaacggcttca cggccatgttgc acctacaatgttgc gatgcacatgttgc 2820
 tag

<210> 32

<211> 941

<212> PRT

<213> fusion protein - human

<400> 32

Met	Leu	Pro	Gly	Leu	Ala	Leu	Leu	Leu	Ala	Ala	Trp	Thr	Ala	Arg	
1				5				10			15				
Ala	Leu	Glu	Val	Pro	Thr	Asp	Gly	Asn	Ala	Gly	Leu	Leu	Ala	Glu	Pro
				20				25			30				

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
 35 40 45
 Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
 50 55 60
 Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
 65 70 75 80
 Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
 85 90 95
 Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
 100 105 110
 Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Ile Ser Asp Ala Leu Leu
 115 120 125
 Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
 130 135 140
 Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
 145 150 155 160
 Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
 165 170 175
 Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
 180 185 190
 Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
 195 200 205
 Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
 210 215 220
 Val Val Glu Val Ala Glu Glu Glu Val Ala Glu Val Glu Glu Glu
 225 230 235 240
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
 245 250 255
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
 260 265 270
 Ala Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Val Arg
 275 280 285
 Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
 290 295 300
 Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
 305 310 315 320
 Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
 325 330 335
 Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
 340 345 350
 Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu
 355 360 365
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
 370 375 380
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Leu Ala Leu Glu Asn
 385 390 395 400
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
 405 410 415
 Asn Met Leu Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
 420 425 430
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
 435 440 445
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
 450 455 460
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
 465 470 475 480
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
 485 490 495
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
 500 505 510
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
 515 520 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
 530 535 540
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
 545 550 555 560
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
 565 570 575
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
 580 585 590
 Glu Val Asn Phe Glu Val Glu Phe Arg His Asp Ser Gly Tyr Glu Val
 595 600 605
 His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
 610 615 620
 Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
 625 630 635 640
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Lys Leu Gly Thr Glu
 645 650 655
 Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Lys Leu
 660 665 670
 Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu Lys Lys Leu
 675 680 685
 Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu Lys Asn Asn
 690 695 700
 Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro Leu Thr Arg
 705 710 715 720
 Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu Glu Gln Leu
 725 730 735
 Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile Leu Lys Met
 740 745 750
 Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu Phe Val Gln
 755 760 765
 Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala Ser Val Glu
 770 775 780
 Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser Ala Thr Ser
 785 790 795 800
 Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu Thr Val Ser
 805 810 815
 Gly Ile Pro Gly Asp Leu Ala Pro Pro Thr Asp Val Ser Leu Gly Asp
 820 825 830
 Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala
 835 840 845
 Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly
 850 855 860
 Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met
 865 870 875 880
 Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp
 885 890 895
 Glu Tyr Gly Gly Asp Ile Gln His Ser Gly Ala Ala Asp Ala Ala
 900 905 910
 Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr
 915 920 925
 Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
 930 935 940

<210> 33
 <211> 63
 <212> PRT
 <213> human

<400> 33
 Glu Glu Ile Ser Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser
 1 5 10 15

Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val
20 25 30
Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val
35 40 45
Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys
50 55 60

<210> 34
<211> 4
<212> PRT
<213> human

<400> 34
Lys Met Asp Ala
1

<210> 35
<211> 7
<212> PRT
<213> human

<400> 35
Gly Val Val Ile Ala Thr Val
1 5

<210> 36
<211> 63
<212> PRT
<213> human

<400> 36
Glu Glu Ile Ser Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser
1 5 10 15
Gly Tyr Glu Val His His Gln Val Leu Val Phe Phe Ala Glu Asp Val
20 25 30
Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val
35 40 45
Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys
50 55 60

<210> 37
<211> 63
<212> PRT
<213> human

<400> 37
Glu Glu Ile Ser Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser
1 5 10 15
Gly Tyr Glu Val His His Gln Val Leu Val Phe Phe Ala Glu Asp Val
20 25 30
Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val
35 40 45
Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys
50 55 60

<210> 38
<211> 4

<212> PRT
<213> human

<400> 38
Asn Leu Asp Ala
1

<210> 39
<211> 63
<212> PRT
<213> human

<400> 39
Glu Glu Ile Ser Glu Val Asn Phe Glu Val Glu Phe Arg His Asp Ser
1 5 10 15
Gly Tyr Glu Val His His Gln Val Leu Val Phe Phe Ala Glu Asp Val
20 25 30
Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val
35 40 45
Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys
50 55 60

<210> 40
<211> 4
<212> PRT
<213> human

<400> 40
Asn Phe Glu Val
1

<210> 41
<211> 5
<212> PRT
<213> human

<400> 41
Val Asn Phe Ala Val
1 5

<210> 42
<211> 5
<212> PRT
<213> human

<400> 42
Val Lys Val Asp Ala
1 5

<210> 43
<211> 5
<212> PRT
<213> human

<400> 43
Trp Lys Met Asp Ala
1 5

<210> 44
<211> 5
<212> PRT
<213> human

<400> 44
Val Lys Ala Asp Ala
1 5

<210> 45
<211> 5
<212> PRT
<213> human

<400> 45
Val Lys Lys Asp Ala
1 5

<210> 46
<211> 5
<212> PRT
<213> human

<400> 46
Val Lys Glu Asp Ala
1 5

<210> 47
<211> 5
<212> PRT
<213> human

<400> 47
Val Lys Ile Asp Ala
1 5

<210> 48
<211> 5
<212> PRT
<213> human

<400> 48
Val Lys Met Ile Ala
1 5

<210> 49
<211> 5
<212> PRT
<213> human

<400> 49
Val Lys Met Asn Ala
1 5

<210> 50
<211> 5
<212> PRT
<213> human

<400> 50
Val Lys Met Glu Ala
1 5

<210> 51
<211> 5
<212> PRT
<213> human

<400> 51
Val Lys Met Asp Glu
1 5

<210> 52
<211> 5
<212> PRT
<213> human

<400> 52
Val Lys Met Asp Lys
1 5

<210> 53
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 53
ggagaggata tcatggagcc agtagat

27

<210> 54
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 54
tacatggccgg ccgcctactt actgctttg

29

<210> 55
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 55
ggatgtgata tc tt ct tc ag cat cac caagg

35

<210> 56
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 56
ctgagatatac aagctactgt cttctatcga acaagc 36

<210> 57
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 57
gcgcgatatac cccaccgtac tcgtcaattc c 31

<210> 58
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 58
ctgctgtggc ggccgcctag ttctgcatct gctc 34

<210> 59
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 59
tgccccgcgc ggccgcgcga tgctgccccgg 30

<210> 60
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 60
atggtgtggc ggccgcagac gccgctgtca cc 32

<210> 61
<211> 36
<212> DNA
<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 61
agcgcacaag cttcccccgcg cagggtcgcg atgctg 36

<210> 62

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 62
ggatgttaagc tttttcttct tcagcatcac caagg 35

<210> 63

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 63
tgcagatatac ctggagccag tagatcctag ac 32

<210> 64

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 64
gctggatatac ctctgctttg atagagaagc 30

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

**(19) World Intellectual Property Organization
International Bureau**



A standard linear barcode is positioned horizontally across the page, consisting of vertical black bars of varying widths on a white background.

(43) International Publication Date
4 September 2003 (04.09.2003)

PCT

(10) International Publication Number
WO 2003/072041 A3

(51) International Patent Classification:

C07H 21/02 (2006.01) **C12N 1/20 (2006.01)**
C12P 21/06 (2006.01) **C12N 15/00 (2006.01)**

ZUCK, Paul, D. [US/US]; 126 East Lincoln Avenue,
Rahway, NJ 07065-0907 (US).

(21) International Application Number:

PCT/US2003/005458

(74) **Common Representative:** MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US)

(22) International Filing Date:

23 February 2003 (23.02.2003)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/360.274 27 February 2002 (27.02.2002) US

(81) **Designated States (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
 - before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:
3 August 2006

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ASSAYS TO MONITOR AMYLOID PRECURSOR PROTEIN PROCESSING

(57) Abstract: The present invention provides DNA constructs, genetically engineered host cells, and methods for identifying inhibitors of amyloid precursor protein (APP) processing. The methods provide for the convenient identification, in a single assay, of inhibitors of β -secretase and γ -secretase as well as other forms of APP processing. The methods rely on fusion proteins of APP and transcription factors in which APP processing releases the transcription factors, allowing the transcription factors to activate transcription of a reporter gene. Inhibitors are identified as substances that block or diminish transcription factor release from the fusion protein, thereby causing a diminution of reporter gene readout.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/05458

A. CLASSIFICATION OF SUBJECT MATTER

IPC: C07H 21/02(2006.01);C12P 21/06(2006.01);C12N 1/20(2006.01),15/00(2006.01)

USPC: 536/23.1;435/69.1,252.3,320.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1; 435/69.1, 252.3, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,811,633 A (Haseltine et al) 1 September 1998 (01.09.1998), nucleic acid encoding HIV-1 Tat transcription factor and trans-activating activity.	1-16
A	US 5,877,015 A (Hardy et al) 2 March 1999 (02.03.1999), SEQ ID NO:4 discloses APP695.	1-16
A	US 5,811,063 (Wadsworth et al) 22 September 1998 (22.09.1998), see claims disclosure of APP695.	1-16

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

20 March 2006 (20.03.2006)

Date of mailing of the international search report

01 JUN 2006

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner of Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Faxsimile No. (571) 273-3201

Authorized officer

Janet Andres

Telephone No. 308-1235

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/05458

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-16

Remark on Protest The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT**BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING**

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group 1 Claim(s) 1-16, drawn to a DNA molecule and method of identifying a substance that inhibits APP processing.

Group 2 Claim(s) 17-20, drawn to a method of identifying a substance that inhibits APP comprising a transcription factor fused to APP.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack unity of invention because they are not so linked as to form a single general inventive concept under PCT Rule 13.1. The species listed do not share a common structure or function.

The species are as follows:

The following claim(s) are generic:

Group 1, Claims 1, 4, 5, 6, 7, 13, 14, 15, 16, all APP species listed therein.

In order for more than one species to be examined, the appropriate additional examination fees must be paid.

Applicant is invited to select one species per generic claim listed per group elected. The inventions listed as Groups 1-2 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Group 1 is drawn to the special technical feature of a DNA molecule comprising APP695 fused to a transcription factor, which is not required by any of the other groups.

Group 2 is drawn to the special technical feature of a fusion protein comprising APP which is cleaved by both beta-secretase and gamma-secretase, which is not required by any of the other groups.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.